

Methods of characterising biological therapeutics

Technical Field

[0001] This disclosure generally relates to methods of characterising samples of biological therapeutics including gene therapy products and other protein-based therapeutic products. It specifically relates to the use of data-independent acquisition mass spectrometry (DIA-MS) for attribute analysis of therapeutic samples.

Background

[0002] This background description is provided for the purpose of generally presenting the context of the disclosure. Contents of this background section are neither expressly nor impliedly admitted as prior art against the present disclosure.

[0003] Production of biological therapeutics like gene therapy products and biologics involve complex bioprocessing workflows which are markedly different from the production of conventional small molecule therapeutics. Product characterisation for these therapeutics needs to take into account new quality attributes including subtle product modifications and trace impurities which can be hard to detect with current analytical methods. For example, gene therapy products, including viral gene therapy products such as recombinant adeno-associated viruses (rAAVs), may contain trace protein contaminants which are not easily identified by routine immunoassays or chromatographic techniques.

[0004] Desired rAAV products are typically composed of an icosahedral capsid assembly of 60 viral proteins (VPs) made up of VP1, VP2, and VP3 in a defined stoichiometry (typically 1:1:10), with a single-stranded DNA molecule encoding the therapeutic gene cassette encased within the rAAV particle. In addition to protecting the therapeutic gene cassette, the rAAV capsid also plays a crucial role in viral infectivity and gene transduction, accentuating the importance of the constituent VPs being well-characterised in gene therapy development. However, sequence homology shared by the VPs and the limited sample availability due to high production costs makes it challenging to use existing analytical methods to determine the concentration, purity and composition of viral gene therapy products. In addition, gene therapy products can also contain a series of process-related impurities as the manufacturing process involves the use of

production and packaging cell lines (e.g. HEK293), helper virus (e.g. Adenovirus), and plasmids. During cell culturing and harvesting, various proteins from multiple sources and species are also employed, including cytokines, growth factors, bovine serum, antibodies, and selection beads (e.g. protein A). Measurement of the level of all these process-related impurities are required based on FDA guidelines to ensure the safety and quality of gene therapy products. However, the mixture of the many components and proteins from different organisms poses a serious challenge to using conventional impurity assays such as enzyme-linked immunosorbent assay (ELISA), which are typically only bind proteins from a single species. As most commercial antibodies used in ELISA are typically raised in goats or rabbits, ELISA tends to suffer from low coverage of human proteins originated from human cell lines used in the manufacturing of gene therapy products.

[0005] Hence, a highly efficient, specific, and sensitive method that can provide reliable analytical results in a timely manner is highly sought after for the characterisation of biological therapeutic products like gene therapy products.

Summary

[0006] In one embodiment, the present disclosure provides a method of characterising a sample of a gene therapy product or a protein-based therapeutic product, the method comprising: performing data-independent acquisition mass spectrometry (DIA-MS) on a sample of a gene therapy product or a protein-based therapeutic product to obtain a spectral signature of the biological sample; and comparing the spectral signature to a spectral signature in a reference spectral library to perform quality attribute analysis of the biological sample.

[0007] The present disclosure also provides a method of preparing a reference spectral library for analysis of DIA-MS spectra, the method comprising performing a plurality of data dependent acquisition mass spectrometry (DDA-MS) on one or more reference samples which have been subjected to a stress condition.

[0008] The present disclosure further provides a method of quality control in production of a gene therapy product or a protein-based therapeutic product, the method comprising: performing attribute analysis of the sample using a method defined herein to obtain attribute analysis results; and comparing the attribute analysis results with predefined attribute reference limits to identify quality control issues in the production of the gene therapy product or protein-based therapeutic product.

Brief Description of the Drawings

[0009] Some embodiments are described, by way of non-limiting example only, with reference to the accompanying drawings in which:

[0010] Figure 1: Schematic diagram of the DIA-MS multi-attribute method (MAM) workflow for streamlined characterisation of gene therapy product critical quality attributes (CQAs) and process-related impurities.

[0011] Figure 2: Peptide mapping of VP1, VP2, VP3 of AAV2 sample using DDA-MS data analysed in PEAKS Studio Xpro.

[0012] Figure 3: Relative quantification of modified peptides in VP1 using DDA-MS data processed in PEAKS Studio Xpro. Note: a = N-terminal acetylation, p = phosphorylation, d = deamidation, o = oxidation and m = methylation.

[0013] Figure 4: Generation of a spectral library for DIA-MS analysis of AAV2 digests.

[0014] Figure 5: Comparison of HCP analysis results from DDA-MS and DIA analysis using both sequence database search (library free) and spectral library search.

[0015] Figure 6: Results from DIA-MS-based MAM for characterisation of viral capsid proteins, PTM analysis and detection of trace host cell protein impurities.

[0016] Figure 7: Diagram comparing conventional methods of characterising gene therapy products and methods according to the embodiments.

[0017] Figure 8: Schematic diagram of a workflow for characterising rAAVs according to the embodiments.

[0018] Figure 9: An example of a DIA-MS based peptide mapping for VP3 of an AAV2.

[0019] Figure 10: Schematic diagram of a rapid streamlined DIA-MS workflow for impurities analysis.

[0020] Figure 11: Spectral analysis results showing that the use of stress-tested reference samples enables confident detection of low-abundance PTMs. (A) DDA-MS spectral analysis of reference samples of monoclonal antibodies (mAbs) stored at 4°C and 37°C for 8 weeks, showing identification and quantification of deamidated amino acid residues. (B) DDA-MS of mAb reference samples exposed to 1% or 3% H₂O₂ for 3

weeks, showing identification and quantification of oxidised amino acid residues. (C) DIA-MS spectral analysis of mAb product samples using stress-tested reference spectra, showing detection of rare deamidated amino acid residues. (D) DIA-MS spectral analysis of mAb product samples using stress-tested reference spectra, showing detection of rare oxidised amino acid residues.

Detailed Description

[0021] This disclosure relates to a streamlined LC-MS (liquid chromatography-mass spectrometry) based workflow for multi-attribute characterisation of samples of protein-containing biological therapeutics, including gene therapy products and protein-based therapeutic products. A spectral library-assisted data-independent acquisition (DIA)-MS and data analysis workflow is used to provide full characterisation of sample proteins, including protein sequence analysis with high sequence coverage, comprehensive post-translational modification (PTM) analysis, determination of protein titres, and quantification of trace host cell and other production-related protein impurities. The disclosed DIA-MS-based multi-attribute method (MAM) requires only minimal sample amounts for comprehensive characterisation of a target therapeutic and unbiased profiling of product- and process-related protein impurities.

[0022] A therapeutic product of this disclosure is any substance intended for therapeutic, preventive, palliative or diagnostic purpose, such as for preventing, diagnosing, monitoring, treating, curing or alleviating any disease, disorder, ailment, injury, handicap or abnormal physical or mental state, or any symptom thereof; for replacing, repairing, regenerating or reconstructing any anatomy; for modifying or replacing any physiological process; or for supporting or sustaining life.

[0023] A biological therapeutic herein is a therapeutic product derived from a biological system, including a whole cell or micro-organism, such as a whole virus or bacterium used as a vaccine; a part of a cell or micro-organism, such as an extracellular vesicle or a subunit vaccine; a plasma-derived product; or a biotechnology-derived substance, such as a recombinant protein or polypeptide, or a recombinant nucleic acid. Protein-containing biological therapeutics include biological therapeutics that contain a native, recombinant and/or modified peptide or polypeptide (such as a peptide or polypeptide that has been conjugated to another substance). For the avoidance of doubt, the peptide or polypeptide need not be the active substance in a protein-containing

biological therapeutic and may be a component of, for example, the formulation or delivery vehicle for the therapeutic.

[0024] A gene therapy product of this disclosure is any therapeutic product whose therapeutic effect is mediated by a recombinant nucleic acid, where the effect of the recombinant nucleic acid relates directly to the recombinant nucleic acid sequence or to the product of the genetic expression of its sequence. The therapeutic effect can be regulation, repair, replacement, addition or deletion of a genetic sequence or modification of genetic material. For example, the gene therapy product may be used for gene supplementation, gene silencing, gene replacement, gene editing, and/or for introducing nucleic acids encoding antigens for the purposes of vaccination. Gene therapy products include but are not limited to viral vectors (e.g., recombinant lentiviruses, adenoviruses and adeno-associated viruses), nucleases and nucleic acids for gene editing (e.g., zinc finger nucleases (ZFNs), transcription activation-like effector nuclease (TALENs), and CRISPR-associated protein 9 (Cas9) with an appropriate guide RNA) and nucleic acid therapeutics (e.g., small interfering RNA (siRNA), microRNA (miRNA), short hairpin RNA (shRNA), messenger RNA (mRNA) and antisense oligonucleotides (ASOs)).

[0025] A protein-based therapeutic product of this disclosure is any therapeutic product whose therapeutic effect is mediated by a polypeptide, preferably by a recombinant polypeptide. Non-limiting examples of protein-based therapeutic products include antibodies, enzymes, hormones, cytokines, growth factors, immune modulators, receptors, signaling proteins, blood factors, antigens, subunit vaccines, and fragments, ligands, inhibitors, agonists and antagonists of such. The protein-based therapeutic product may contain a modified polypeptide, such as a polypeptide that has been conjugated to another drug substance or to a label for imaging.

[0026] A product attribute is a feature or characteristic of a therapeutic product which may affect drug safety and/or efficacy. For protein-based or protein-containing therapeutics, such attributes can include amino acid sequence variations (e.g., substitutions, insertions, deletions), glycosylation patterns and other post-translational modifications (PTMs) to amino acid side chains, and the presence of protein or peptide impurities from cell culture or downstream purification processes. Critical quality attributes (CQAs) are those attributes which have a direct impact on drug safety and/or efficacy, thus CQAs often form part of the regulatory assessment for particular therapeutics or classes of therapeutics.

[0027] While conventional data-dependent acquisition (DDA)-MS methods are useful for characterising sequence identity and post-translational modifications of sample proteins, this method faces limitations in profiling host-cell proteins (HCP) or trace protein contaminants in protein samples due to the stochastic nature of its precursor selection methodology and the low sampling efficiency of current mass spectrometers. For example, the typical DDA method only selects the top 'N' precursors in descending order of magnitude and is thus biased towards more abundant peptides, leading to relatively lower-abundance peptides (such as from trace contaminants) being missed in data analysis.

[0028] The DIA-MS method of this disclosure is advantageous in its unbiased approach to peptide profiling. In DIA-MS-based methods, tandem mass spectra are acquired by fragmenting all ions that enter the mass spectrometer at a given time (DIA) or by sequentially focusing on a narrow m/z window of precursors and fragmenting all precursors detected within that window (SWATH-DIA). This allows the acquisition of a full proteomic profile of precursors and product-ions in a single analysis, offering highly sensitive detection capabilities and generating comprehensive data sets for use in quantitative analysis without being limited to profiling predefined peptides of interest. Results of a DIA-MS analysis are presented in the form of a spectral signature or a plurality of spectral signatures generated during the DIA-MS operations. DIA-MS advantageously allows for profiling of trace protein impurities whose concentrations may be too low to be picked up by DDA-MS methods.

[0029] Another advantage of the described workflow is that it is applicable not only to purified samples of target therapeutics (e.g., viral vectors) but also to samples of varying purity, including semi-purified products, stressed samples used for stability testing, and samples from various intermediate upstream and downstream processes. Furthermore, the disclosed workflow is more readily employable for process optimisation during therapeutic manufacturing compared to other protein analytical methods like antibody-based immunoassays. For example, the methods herein can be integrated into an automated workflow for in-process monitoring of multiple quality attributes during therapeutic manufacturing. The methods can also be used to guide the removal of specific protein impurities as well as to guide the development of immunoassays for these impurities, thus complementing existing analytical methods.

[0030] To perform high-quality quantitative analysis using DIA-MS data, a specific spectral library is preferred over the use of a computational prediction protocol based on a protein sequence database (library-free analysis). The methods of the present disclosure also include generation of a comprehensive spectral library to accommodate the DIA strategy for characterising trace protein impurities. Advantageously, this spectral library merges spectra from the target therapeutic and the proteome of the host cell for accurate and reproducible identification and quantification of trace host cell protein (HCP) contaminants. In addition, the spectral library can be built using reference samples which have been stressed-tested under conditions capable of inducing a chemical or physical change to a target therapeutic (e.g., a change to an amino acid functional group, or aggregation of the therapeutic). This provides reference spectra which can be used to identify rare modifications in the therapeutic product, including post-translational modifications and degraded products.

[0031] Exemplary, non-limiting embodiments of a method of characterising a sample of a gene therapy product or a protein-based therapeutic product using a DIA-MS-based multiple attribute method (MAM), will now be disclosed.

[0032] This disclosure provides a method of characterising a sample of a gene therapy product or a protein-based therapeutic product, the method comprising: performing data-independent acquisition mass spectrometry (DIA-MS) on a sample of the gene therapy product or protein-based therapeutic product to obtain a spectral signature of the sample, and comparing the spectral signature to a spectral signature in a reference spectral library to perform attribute analysis of the sample.

[0033] In an embodiment, attribute analysis in the method includes at least one of: characterisation of amino acid sequences of peptides and proteins in the sample, post-translational modification (PTM) analysis of the proteins in the biological sample, protein impurity analysis of the sample, or determination of the titre of the gene therapy product or protein-based therapeutic product in the sample.

[0034] In an embodiment, attribute analysis in the method includes all of: characterisation of amino acid sequences of peptides and proteins in the sample, post-translational modification (PTM) analysis of the proteins in the biological sample, protein impurity analysis of the sample, or determination of the titre of the gene therapy product or protein-based therapeutic product in the sample.

[0035] Advantageously, the method provides a comprehensive data acquisition and analytics workflow to concurrently assess multiple critical attributes of a target therapeutic, especially the identification and quantification of trace product- and process-related protein impurities and rare PTMs.

[0036] In an embodiment, characterisation of amino acid sequence is performed using peptide mapping, wherein the mass spectrometric spectra of peptides derived from enzymatic or chemical fragmentation of the protein are compared to peptide spectra previously determined for known proteins to identify sequence deviations.

[0037] In an embodiment, the PTM analysis includes identification of one or more of amino acid acetylation, amino acid phosphorylation, amino acid deamidation, amino acid oxidation or amino acid methylation. Other types of post-translational modifications contemplated include hydroxylation, glycation, reduction, isomerisation, glycosylation, lipidation, ubiquitination, nitrosylation and cyclisation at one or more amino acid residues, as well as peptide or protein cyclisation, proteolysis or fragmentation.

[0038] In an embodiment, protein impurity analysis comprises identification and/or quantification of one or more of a host cell protein, a protein from a culture medium, a protein contaminant from a production process for the gene therapy product or polypeptide-based therapeutic product, or a protein, protein fragment or peptide which is determined not to be a component of the reference spectral library.

[0039] The host cell may be any cell or cell line used in the production of a target gene therapy product or protein-based therapeutic product, including but not limited to a prokaryotic cell, a eukaryotic cell, a bacterial cell such as an *E. coli* cell, an insect cell, a fungal cell such as a yeast cell, an avian cell, and a mammalian cell such as a Chinese hamster ovary (CHO) cell or a human embryonic kidney (HEK) cell. The host cell protein may be any endogenous or exogenous protein or protein fragment derived from the host cell, but excludes the biological therapeutic produced by the host cell.

[0040] The culture medium protein may be any protein present in the medium or secreted into the medium by the host cell during therapeutic manufacturing, but excludes the biological therapeutic produced by the host cell. Such contaminants may include, for example, cytokines, growth factors, serum components, antibodies, proteins from viral contaminants, and proteins from helper viruses.

[0041] The production process for a gene therapy product or protein-based therapeutic product includes all upstream and downstream bioprocessing workflows, including but not limited to cell culture, cell harvest, product isolation, product purification, product formulation and product packaging. Protein contaminants from production processes include any protein introduced as part of a production process but which is not intended for inclusion in the packaged therapeutic product. For example, protein contaminants from purification processes can include proteins used in affinity chromatography, e.g., Protein A, Protein G.

[0042] Protein, protein fragments and peptides which are determined not to be a component of the reference spectral library include all other functional and non-functional proteins or peptides without representative spectra in the reference spectral library, including, for example, protein mutants, fusion proteins, degraded proteins or peptides, and other forms of post-translational modifications or chemical changes which may give rise to spectral variabilities that deviate from the reference spectra

[0043] In an embodiment, the gene therapy product is a recombinant viral vector or a nucleic acid therapeutic (NAT). Non-limiting examples of recombinant viral vectors include retroviruses, lentiviruses, adenoviruses, adeno-associated viruses and alphaviruses, including replication-competent and replication-defective forms thereof. The viral vectors may be used for gene supplementation, gene silencing, gene replacement, gene editing, and/or for introducing nucleic acids encoding antigens for the purpose of vaccination.

[0044] Nucleic acid therapeutics include polynucleotides (e.g., single-stranded or double-stranded DNA or RNA) which are not packaged in viral vectors. The nucleic acid therapeutic may be, for example, mRNA, antisense oligonucleotides (ASOs), and polynucleotides mediating RNA interference, such as small interfering RNA (siRNA), microRNA (miRNA), short hairpin RNA (shRNA). Proteins used during the manufacture of NATs (such as enzymes used for *in vitro* transcription or capping of mRNA or in downstream purification processes) may present as impurities in the final therapeutic product and may be detected using methods of the present disclosure.

[0045] In an embodiment, the gene therapy product is a recombinant adeno-associated virus (rAAV). rAAVs generally comprise a viral capsid encapsulating a single-stranded polynucleotide. The viral capsid may comprise capsid proteins VP1, VP2 and/or VP3.

[0046] In one embodiment, the protein-based therapeutic product is a monoclonal antibody (e.g., IgA, IgD, IgE, IgG, IgM, IgNAR or camelid antibody) or an antigen-binding fragment thereof (e.g., Fab, Fv, vNAR, nanobody, etc.), a monoclonal antibody derivative (e.g., a single chain Fv (scFv), a single chain Fab (scFab), a minibody, diabody, triabody, tetrabody, bispecific antibody, trispecific antibody, etc.), a protein drug or vaccine (e.g., a hormone, enzyme, cytokine, growth factor, immune modulator, receptor, signaling protein, blood factor, antigen, ligand, agonist, antagonist or inhibitor), or an antibody-drug conjugate.

[0047] In an embodiment, the attribute analysis is performed on viral capsid proteins present in the sample. The viral capsid proteins may be, for example, VP1, VP2 and/or VP3.

[0048] In an embodiment, the method comprises digesting the sample, e.g., by enzymatic or chemical proteolysis, before performing the DIA-MS on the sample.

[0049] In an embodiment, the reference spectral library is prepared by performing a plurality of data dependent acquisition mass spectrometry (DDA-MS) on one or more reference samples. A reference sample is a sample used to generate a reference spectral signature against which a product spectral signature may be compared to perform attribute analysis of the sample.

[0050] The reference sample may be a pure sample of the gene therapy product or protein-based therapeutic product or a sample from any stage of a production process for the therapeutic. Where a host cell is used to produce the therapeutic, the reference sample preferably includes a sample of the host cell, e.g., a cell culture sample of the host cell. Reference samples also include, but are not limited to, protein-containing chromatography resins (e.g., Protein A or Protein G resins), eluents and diluents, and enzymes for *in vitro* modification of a therapeutic (e.g., capping enzymes for mRNA). In an embodiment, the reference sample has been subjected to a stress condition, i.e., the reference sample has been stress-tested. A stress condition according to this disclosure is a condition capable of inducing a chemical and/or physical change in a therapeutic product. For example, the stress condition may lead to the formation of an attribute at an amino acid of the therapeutic product. A stress condition may be due to a production and/or storage process of the therapeutic. Alternatively, a stress condition may be due to deviations from typical production and/or storage conditions. Non-limiting examples of stress conditions include specific conditions of or changes in pH, temperature, osmolality,

light (e.g., intensity, wavelength), pressure, solvent, salinity, oxygen content, agitation, formulation, cell line or culture medium, or exposure to chemicals and materials (e.g., detergents, preservatives, oxidising or reducing agents, organic solvents, plastics, metals, chaotropic salts, etc.). Stress testing has the potential to reveal rare attributes that may arise during production and/or storage of the therapeutic product. The use of stress-tested reference samples is thus advantageous in building a more comprehensive reference spectral library for identifying rare attributes in product samples.

[0051] In one embodiment, both stress-tested and non-stress-tested reference samples are used to prepare the reference spectral library.

[0052] In one embodiment, the product sample is subjected or is one that has been subjected to a stress condition. Product samples may be stress-tested to identify potential physical or chemical changes that may arise during production and/or storage of the therapeutic product. Both stress-tested and non-stress-tested product samples may be used in the methods of the present disclosure.

[0053] The product sample and reference sample may be subjected to the same stress conditions or to different stress conditions.

[0054] In one embodiment, the stress condition is an exposure to one or more of: a change in temperature, a change in pH, a change in osmolality, a change in salinity, a freeze/thaw cycle, an oxidising environment, a reducing environment, ultraviolet (UV) light, or long-term storage.

[0055] Also provided herein is a method of determining whether a sample of a gene therapy product or a protein-based therapeutic product has been subjected to a stress condition. The method may comprise performing data-independent acquisition mass spectrometry (DIA-MS) on the sample to obtain a spectral signature of the sample; and comparing the spectral signature to a spectral signature in a reference spectral library to perform attribute analysis of the sample. The method may comprise identifying a sample that has been subjected to a stress condition based on a similar or identical spectral signature in the reference spectral library. Alternatively, the method may comprise identifying a sample that has not been subjected to a stress condition based on a similar or identical spectral signature in the reference spectral library.

[0056] The present disclosure also provides a method of preparing a reference spectral library for analysis of DIA-MS spectra, the method comprising performing a

plurality of data dependent acquisition mass spectrometry (DDA-MS) on one or more reference samples which have been subjected to a stress condition. Reference samples and stress conditions are as defined in this disclosure. The use of stressed reference samples has the advantage of providing reference spectra for new attributes that may arise from production and/or storage processes for the therapeutic product, thus enabling more efficient and comprehensive attribute analysis (e.g., post-translational modification analysis) of product samples.

[0057] A combination of stress-tested and non-stress-tested reference samples may be used to prepare the reference spectral library.

[0058] The present disclosure further provides a method of quality control in production of a gene therapy product or a protein-based therapeutic product, the method comprising performing attribute analysis of the sample using the method defined herein to obtain attribute analysis results, and comparing the attribute analysis results with predefined attribute reference limits to identify a quality control issue in the production of the gene therapy product or protein-based therapeutic product. The method may further comprise obtaining a sample for attribute analysis.

[0059] An attribute reference limit typically reflects the extent to which a change in an attribute is due to natural variation of a production or storage process. The attribute reference limit can be established based on prior experimentation and/or stress testing where normal (acceptable) and unacceptable values for an attribute is acquired. The reference limit can be the level of an attribute, a percentage change in the level of an attribute, or the presence or absence of a spectral feature (e.g., a spectral peak, component, ion, ratio, etc.). Accordingly, a quality control issue can arise if the attribute analysis results exceeds a reference limit (e.g., for the concentration of an impurity), is below a reference limit (e.g., for the occurrence of a desired post-translational modification) or deviates from a reference spectral signature.

[0060] In one embodiment, the quality control issue comprises one or more of: amino acid sequence modifications in the gene therapy product or protein-based therapeutic product; aberrant post-translational modification (PTM) of the gene therapy product or protein-based therapeutic product; or presence of protein impurities in the sample.

[0061] Preferably, the amino acid sequence modification is one that affects or has the potential to affect a physical or biological property of a protein in the gene therapy product

or protein-based therapeutic product, e.g., the modification may change the folding of a protein or a biological activity of the protein. Exemplary sequence modifications include but are not limited to non-conservative substitution, addition and/or deletion of one or more amino acid residues, inversion of a stretch of amino acids, and sequence variations arising from atypical exon splicing.

[0062] Aberrant PTMs include non-native amino acid acetylation, phosphorylation, deamidation, oxidation, methylation, hydroxylation, glycation, reduction, isomerisation, glycosylation, lipidation, ubiquitination, nitrosylation and/or cyclisation, as well as non-native peptide or protein cyclisation, proteolysis and/or fragmentation.

[0063] Protein impurities may derive from any part of the production process for the gene therapy product or protein-based therapeutic product. Non-limiting examples of protein impurities include host cell proteins, proteins from culture media and proteins from purification processes.

[0064] The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

[0065] Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

[0066] Throughout this specification and the claims which follow, unless the context requires otherwise, the phrase "consisting essentially of", and variations such as "consists essentially of" will be understood to indicate that the recited element(s) is/are essential i.e. necessary elements of the invention. The phrase allows for the presence of other non-recited elements which do not materially affect the characteristics of the invention but excludes additional unspecified elements which would affect the basic and novel characteristics of the method defined.

[0067] As used in this specification, “and/or” refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

[0068] As used in this application, the singular form “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “an agent” includes a plurality of agents, including mixtures thereof.

[0069] The scope of this disclosure encompasses all changes, substitutions, variations, alterations, and modifications to the example embodiments described or illustrated herein that a person having ordinary skill in the art would comprehend. The scope of this disclosure is not limited to the example embodiments described or illustrated herein. Moreover, although this disclosure describes and illustrates respective embodiments herein as including particular components, elements, feature, functions, operations, or steps, any of these embodiments may include any combination or permutation of any of the components, elements, features, functions, operations, or steps described or illustrated anywhere herein that a person having ordinary skill in the art would comprehend. Although this disclosure describes or illustrates particular embodiments as providing particular advantages, particular embodiments may provide none, some, or all of these advantages.

[0070] Unless otherwise defined, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

[0071] Certain embodiments of the invention will now be described with reference to the following examples which are intended for the purpose of illustration only and are not intended to limit the scope of the generality hereinbefore described.

EXAMPLES

[0072] The following embodiment is an LCMS-based multi-attribute method (MAM) for multi-attribute characterisation of gene therapy products such as a recombinant adeno-associated virus (rAAV). The multi-attribute method contains two main sections: (a) characterisation of viral capsid proteins including their protein sequences and post-translational modifications, and b) characterisation of trace protein impurities originating from product manufacturing processes.

(a) Characterisation of rAAV capsid proteins

Preparation of rAAV samples for digestion and peptide mapping

[0073] One of the key challenges in rAAV product characterisation is the limitation of samples available for method development and optimization due to high production costs. The present disclosure provides an efficient digestion workflow with minimal sample requirement and is applicable to samples with varying protein concentration and buffer conditions. One to three micrograms of rAAV samples is sufficient for sample preparation. In some embodiments, S-TRAP™ (Protifi) was used for enzymatic (i.e. trypsin, pepsin) digestion before LC-MS according to manufacture instructions Both DDA-MS and DIA-MS methods were used for the analysis of rAAV digest samples for comparison.

Peptide mapping and post-translational modification analysis

[0074] Using the DDA-MS method, more than 91% sequence coverage was achieved for all 3 capsid proteins (VP1: 91%, VP2: 92%, VP3: 93%) with a digestion process taking less than 1.5 hours to complete. Data analysis may be done by using PEAKS® Studio Xpro (Bioinformatics Solutions Inc., Canada) with FDR of 0.1% for peptide match and Ascore of 20 for PTM analysis (p-value <0.001). The analysis also enabled simultaneous PTM analysis for critical modifications such as N-terminal acetylation, deamidation, and oxidation. To demonstrate the capability of the method in detecting other PTMs, the data was further analyzed using an exhaustive workflow for complete PTM analysis. Figure 3 shows relative PTM quantification of modified amino acid residues in VP1, which include selected modifications known to be critical such as N-terminal acetylation, phosphorylation, deamidation, oxidation, and methylation.

(b) Characterisation of trace host cell protein impurities

Generation of in-house HEK293-AAV2 spectral library

[0075] Using an rAAV production cell line, HEK293, a spectral library was constructed by acquiring DDA-MS data of 78 high-pH fractionated samples. The raw DDA data files were processed using the ProteinPilot™ software (V5.0.2 (SCIEX, Macclesfield, UK) and matched against a human UniProt Proteome reference database (downloaded in September 2021 from uniprot.org/taxonomy/9606; Human_uniprot-taxonomy_9606.faa). The human RefSeq assembly was appended with the Biognosys iRT fusion protein sequence. The search result files were further used as the input libraries for spectral

library generation using Spectronaut™ software (Version: 15.7.220308.50606, Rubin). A secondary spectral library was also generated by directly loading the DDA data files on Spectronaut which uses the integrated database search engine called Pulsar for Spectral library generation. To generate a deep, precise, and robust library, different search archives (Protein Pilot and Pulsar-based DDA search results) were tested to generate a more comprehensive spectral library. In this new library, any newly identified peptides were added to the existing proteins and any newly detected proteins are added if not present in the base library. DDA data generated from peptide mapping of purified rAAV digest was also processed using the aforementioned workflow to generate a spectral library for AAV2. By merging both the AAV2 and HEK293 spectral libraries together, a HEK293-AAV2 spectral library was generated for processing of DIA-MS data generated from the analysis of rAAV sample digests (Figure 4). Compared with DDA-MS and DIA-MS-based methods with sequence database search (library free), the DIA-MS-based multi-attribute method, together with the in-house HEK293-AAV2 spectral library, introduced by the current disclosure presents a more superior capability with higher HCP coverage (Figure 5).

Using HEK293-AAV2 spectral library for multi-attribute monitoring of AAV2 samples

[0076] Some embodiments used a HEK293-AAV2 spectral library in library-based processing of DIA-MS data for multi-attribute monitoring of rAAV2 samples using Spectronaut™ software (Biognosys, Switzerland). This method allows the simultaneous characterisation of the HCP profile, rAAV protein sequence identification, and PTM quantification (Figure 6).

[0077] Technical features of some embodiments include:

- A multi-attribute method (MAM) for characterisation and quantification of biological therapeutic products, especially gene therapy products like viral vectors, and product- and process-related impurities using a single unbiased quantitative DIA-MS workflow.
- A peptide mapping workflow that provides a comprehensive characterisation of biological therapeutic products, especially gene therapy products like viral vectors, including critical quality attributes such as protein sequence, PTM analysis, and individual host cell protein impurity, all within a few hours from sample to results.

- A comprehensive spectral library that includes the host cell proteome and target therapeutic for accurate and sensitive identification and quantification of trace protein contaminants (e.g., host cell proteins) in biological therapeutic products (process intermediates and final purified products).
- An automated in-process monitoring workflow for multiple quality attributes of biological therapeutics (especially gene therapy products like viral vectors), serving as an alternative orthogonal workflow to antibody-based immunoassays for determination of titres and quantification of host cell proteins (HCPs).

Claims

1. A method of characterising a sample of a gene therapy product or a protein-based therapeutic product, the method comprising:
 - performing data-independent acquisition mass spectrometry (DIA-MS) on the sample to obtain a spectral signature of the sample; and
 - comparing the spectral signature to a spectral signature in a reference spectral library to perform attribute analysis of the sample.
2. The method of claim 1, wherein the attribute analysis comprises at least one of:
 - characterisation of amino acid sequences of peptides and proteins in the sample;
 - post-translational modification (PTM) analysis of the proteins in the sample;
 - protein impurity analysis of the sample; or
 - determination of the titre of the gene therapy product or protein-based therapeutic product in the sample.
3. The method of claim 2, wherein the attribute analysis comprises:
 - characterisation of amino acid sequences of peptides and proteins in the sample;
 - post-translational modification (PTM) analysis of the proteins in the sample;
 - protein impurity analysis of the sample; and
 - determination of the titre of the gene therapy product or protein-based therapeutic product in the sample.
4. The method of claim 2 or 3, wherein the PTM analysis includes identification of one or more of: amino acid acetylation, amino acid phosphorylation, amino acid deamidation, amino acid oxidation or amino acid methylation.
5. The method of claim 2 or 3, wherein protein impurity analysis comprises identification and/or quantification of one or more of: a host cell protein, a protein from a culture medium, a protein contaminant from a production process for the gene therapy product or protein-based therapeutic product, or a protein, protein

fragment or peptide which is determined not to be a component of the reference spectral library.

6. The method of any one of claims 1 to 5, wherein the gene therapy product is a recombinant viral vector or a nucleic acid therapeutic.
7. The method of any one of claims 1 to 5, wherein the protein-based therapeutic product is a monoclonal antibody or an antigen-binding fragment thereof, a monoclonal antibody derivative, a protein drug or vaccine, or an antibody-drug conjugate.
8. The method of claim 6, wherein the gene therapy product is a recombinant adeno-associated virus (rAAV)
9. The method of claim 8, wherein attribute analysis is performed on viral capsid proteins in the rAAV.
10. The method of claim 6, wherein the gene therapy product is an RNA therapeutic.
11. The method of any one of claims 1 to 10, wherein the reference spectral library is prepared by performing a plurality of data dependent acquisition mass spectrometry (DDA-MS) on one or more reference samples
12. The method of claim 11, wherein the reference sample has been subjected to a stress condition.
13. The method of any one of claims 1 to 12, wherein the product sample has been subjected to a stress condition.
14. The method of claim 12 or 13, wherein the stress condition is an exposure to one or more of: a change in temperature, a change in pH, a change in osmolality, a change in salinity, a freeze/thaw cycle, an oxidising environment, a reducing environment, ultraviolet (UV) light, or long-term storage.
15. A method of preparing a reference spectral library for analysis of DIA-MS spectra, the method comprising performing a plurality of data dependent acquisition mass

spectrometry (DDA-MS) on one or more reference samples which have been subjected to a stress condition.

16. A method of quality control in production of a gene therapy product or a protein-based therapeutic product, the method comprising:

performing attribute analysis of the sample using the method of any one of claims 1 to 10 to obtain attribute analysis results; and

comparing the attribute analysis results with predefined attribute reference limits to identify a quality control issue in the production of the gene therapy product or protein-based therapeutic product.

17. The method of claim 16, wherein the quality control issue comprises one or more of:

amino acid sequence modifications in the gene therapy product or protein-based therapeutic product;

aberrant post-translational modification (PTM) of the gene therapy product or protein-based therapeutic product; or

presence of protein impurities in the sample.