



(51) International Patent Classification:

G01N 33/68 (2006.01) G16C 20/20 (2019.01)

G01N 21/65 (2006.01) C07K 14/00 (2006.01)

G16B 15/00 (2019.01)

(21) International Application Number:

PCT/US2022/021289

(22) International Filing Date:

22 March 2022 (22.03.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/165,048 23 March 2021 (23.03.2021) US

63/165,057 23 March 2021 (23.03.2021) US

63/165,055 23 March 2021 (23.03.2021) US

63/165,040 23 March 2021 (23.03.2021) US

(71) Applicant: JANSSEN BIOTECH, INC. [US/US];

800/850 Ridgeview Drive, Horsham, PA 19044 (US).

(72) Inventors: O'MAHONY-HARTNETT, Caitlin; Barnahely, Ringaskiddy, Co. Cork, P43 FA46 (IE). MC-CARTHY, Barry J.; Barnahely, Ringaskiddy, Co. Cork, P43 FA46 (IE). RODE, Christopher W.; 1400 McKean Road, Spring House, PA 19477 (US). HAYES, Ronan; Barnahely, Ringaskiddy, Co. Cork, P43 FA46 (IE). MADDEN, Fiona; Barnahely, Ringaskiddy, Co. Cork, P43 FA46 (IE). MO, Jingjie; 1400 McKean Road, Spring House, PA 19477 (US). MASLANKA, Francis C.; 1400 McKean Road, Spring House, PA 19477 (US). CLARK, Kevin; 1400 McKean Road, Spring House, PA 19477 (US). TROUT, Daniel A.; 1400 McKean Road, Spring House, PA 19477 (US). RANA, Pushpraj; 1400 McKean Road,

Spring House, PA 19477 (US). GRATSLIA, Emmanouela; Einsteinweg 101, 2333 CB Leiden (NL). RAFFERTY, Carl; Barnahely, Ringaskiddy, Co. Cork, P43 FA46 (IE). BALSS, Karin M.; 1400 McKean Road, Spring House, PA 19477 (US).

(74) Agent: WEISSER, Tamera M. et al.; Jones Day, 250 Vesey Street, New York, NY 10281-1047 (US).

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

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

(54) Title: MULTIPARAMETER MATERIALS, METHODS AND SYSTEMS FOR ENHANCED BIOREACTOR MANUFACTURE

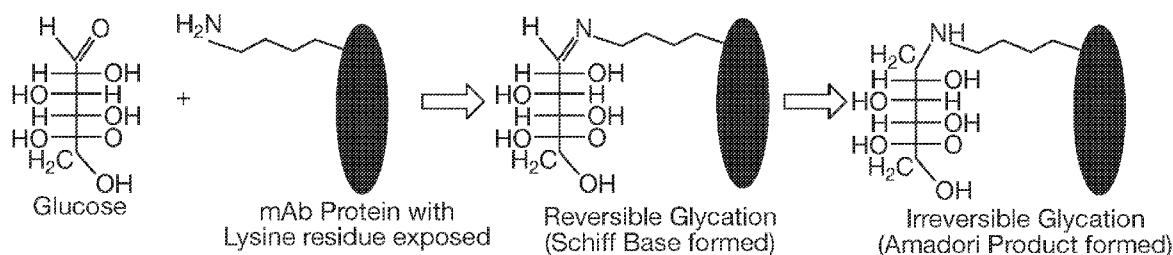


FIG. 1A

(57) Abstract: Methods for determining glycation on a molecule and/or a glycan structure on a glycosylated molecule through the use of a combination of spectroscopic analysis and chemometric modeling are described. In addition, methods and systems for producing a molecule with a desired level of glycation, including a non-glycated molecule, and/or a desired level of a glycan structure on a glycosylated molecule are described.

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

**MULTIPARAMETER MATERIALS, METHODS AND SYSTEMS FOR ENHANCED
BIOREACTOR MANUFACTURE**

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Serial No. 63/165,040 filed March 23, 2021,
5 U.S. Serial No. 63/165,048 filed March 23, 2021, U.S. Serial No. 63/165,055 filed March 23,
2021, and U.S. Serial No. 63/165,057 filed March 23, 2021, the disclosure of each of which is
incorporated by reference herein in its entirety.

1. FIELD

[0001] The present disclosure relates, in part, to multiparameter materials, methods and
10 systems for enhanced bioreactor manufacture. Specifically, the disclosure relates to methods and
systems to control and monitor glycosylation of therapeutic proteins [*e.g.*, recombinant proteins
and/or monoclonal antibodies, (mAbs)] during the production process.

2. BACKGROUND

[0002] Glycation and glycosylation have been deemed a Critical Quality Attribute (CQA) to
15 be considered in the production of therapeutic proteins. For example, glycation can potentially
affect bioactivity and molecular stability of therapeutic proteins. Additionally, glycosylation of
therapeutic proteins can influence the proteins' aggregation, solubility and stability both *in vitro*
and *in vivo*. Therefore, detection and characterization of glycation and/or glycosylation are an
important aspect of the production of therapeutic proteins.

20 **3. SUMMARY**

[0003] Against this background, the inventors of the present disclosure discovered methods
and systems to control and monitor glycation and/or glycosylation of molecules (such as a
therapeutic protein disclosed in Section 5.1) during the production process.

[0004] In one aspect, provided herein is a method for determining a glycan structure on a
25 glycosylated molecule, the method comprising: obtaining, for each of a plurality of runs, levels
for one or more glycan structures on the glycosylated molecule using a process analytical
technology (PAT) tool, wherein the obtaining is within one or more first bioreactors having a
first volume equal to or below a first threshold, the PAT tool obtaining spectral data; generating

one or more regression models based on the obtained spectral data that correlate levels of the one or more glycan structures on the glycosylated molecule with the obtained spectral data; measuring the one or more glycan structures on the glycosylated molecule using the PAT tool, wherein the measuring is within one or more second bioreactors having a second volume equal to or above a second threshold to result in measured spectral data; and determining, by at least one computing device using the generated one or more regression models and based on the measured spectral data, levels of the one or more glycan structures on the glycosylated molecule within one or more second bioreactors.

[0005] In one embodiment, the method further includes refining the one or more regression models based on the combination of the obtained spectral data and the measured spectral data.

[0006] In one embodiment, the method further includes maintaining one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule.

[0007] In one embodiment, the method further includes selectively modifying one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule. In one embodiment, the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof. In one embodiment, the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. In one embodiment, the concentration of glucose is automatically modified based on the measured spectral data.

[0008] In one embodiment, the method further includes purifying the glycosylated molecule.

[0009] In one embodiment, the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[0010] In one embodiment, the obtaining is within two or more bioreactors having different volumes. In one embodiment, the first threshold is about 250 liters or less. In one embodiment, the first threshold is about 100 liters or less. In one embodiment, the first threshold is about 50 liters or less. In one embodiment, the first threshold is about 25 liters or less. In one embodiment,

the first threshold is about 10 liters or less. In one embodiment, the first threshold is about 5 liters or less. In one embodiment, the first threshold is about 2 liters or less. In one embodiment, the first threshold is about 1 liter or less. In one embodiment, the second threshold is about 1,000 liters or more. In one embodiment, the second threshold is about 2,000 liters or more. In one
5 embodiment, the second threshold is about 5,000 liters or more. In one embodiment, the second threshold is about 10,000 liters or more. In one embodiment, the second threshold is about 15,000 liters or more. In one embodiment, the second threshold is about 10,000 to about 25,000 liters. In one embodiment, the second threshold is about 15,000 liters. In one embodiment, the second threshold is at least 5X greater than the first threshold. In one embodiment, the second
10 threshold is at least 10X greater than the first threshold. In one embodiment, the second threshold is at least 100X greater than the first threshold. In one embodiment, the second threshold is at least 500X greater than the first threshold. In one embodiment, the first volume is about 0.5 to about 250 liters. In one embodiment, the first volume is about 1 to about 50 liters. In one embodiment, the first volume is about 1 to about 25 liters. In one embodiment, the first volume is
15 about 1 to about 10 liters. In one embodiment, the first volume is about 1 to about 5 liters. In one embodiment, the second volume is about 1,000 to about 25,000 liters. In one embodiment, the second volume is about 2,000 to about 25,000 liters. In one embodiment, the second volume is about 5,000 to about 25,000 liters. In one embodiment, the second volume is about 10,000 to about 25,000 liters. In one embodiment, the second volume is about 15,000 to about 25,000
20 liters.

[0011] In one embodiment, the PAT tool utilizes or otherwise comprises Raman spectroscopy.

[0012] In one embodiment, the one or more regression models comprise a partial least squares (PLS) model.

25 **[0013]** In one embodiment, the glycosylated molecule comprises a monoclonal antibody (mAb). In one embodiment, the glycosylated molecule comprises a non-mAb.

[0014] In one embodiment, the determining step is performed on-site. In one embodiment, the determining step is performed off-site.

[0015] In one embodiment, the determining step is performed in-line, at-line, on-line, off-
30 line, or a combination thereof. In one embodiment, the determining step is performed on-line. In

one embodiment, the determining step is performed at-line. In one embodiment, the determining step is performed off-line. In one embodiment, the determining step is performed in-line.

[0016] In one embodiment, the obtaining includes receiving data characterizing the spectral data from the PAT tool.

5 **[0017]** In one embodiment, the generating is executed by one or more computing devices.

[0018] In one aspect, provided herein is a method of producing a glycosylated molecule having a desired glycan structure, the method comprising: measuring one or more glycan structures using a process analytical technology (PAT) tool to result in spectral data, wherein the measuring is within a bioreactor having a volume of equal to or above 1,000 liters; determining,
10 by at least one computing device using one or more regression models and based on the measured spectral data, a level of the one or more glycan structures within the bioreactor, wherein the one or more regression models are generated using trial runs from at least one bioreactor having a volume less than or equal to 50 liters and at least one bioreactor having a volume equal to or above 1,000 liters; and maintaining one or more operating parameters of the
15 bioreactor when: the level of a desired glycan structure is above a pre-defined threshold, or the level of an undesired glycan structure is below a pre-defined threshold; or selectively modifying one or more operating parameters of the bioreactor when: the level of a desired glycan structures is below a pre-defined threshold, or the level of an undesired glycan structure is above a pre-defined threshold.

20 **[0019]** In one embodiment, the measuring is performed in-line, at-line, on-line, off-line, or a combination thereof. In one embodiment, the measuring is performed in-line. In one embodiment, the measuring is performed at-line. In one embodiment, the measuring is performed off-line. In one embodiment, the measuring is performed on-line.

[0020] In one embodiment, the measuring occurs more than once daily. In one embodiment,
25 the measuring occurs about every 5 to 60 minutes. In one embodiment, the measuring occurs about every 10 to 30 minutes. In one embodiment, the measuring occurs about every 10 to 20 minutes. In one embodiment, the measuring occurs about every 12.5 minutes.

[0021] In one embodiment, the bioreactor volume is about 2,000 liters or more. In one embodiment, the bioreactor volume is about 5,000 liters or more. In one embodiment, the

bioreactor volume is about 10,000 liters or more. In one embodiment, the bioreactor volume is about 15,000 liters or more. In one embodiment, the bioreactor volume is about 10,000 liters to about 25,000 liters. In one embodiment, the bioreactor volume is about 15,000 liters.

[0022] In one embodiment, the determining step is performed on-site. In one embodiment, the determining step is performed off-site.

[0023] In one embodiment, the desired glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[0024] In one embodiment, the bioreactor is a batch reactor, a fed-batch reactor, a perfusion reactor, or a combination thereof.

[0025] In one embodiment, the one or more operating parameters include a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof. In one embodiment, the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. In one embodiment, the concentration of glucose is automatically modified based on the measured spectral data.

[0026] In one embodiment, the PAT tool includes Raman spectroscopy.

[0027] In one embodiment, the one or more regression models include a partial least squares (PLS) model.

[0028] In one aspect, provided herein is a system for producing one or more glycosylated molecules and the system includes a means for culturing a glycosylated molecule-producing cell line; means for measuring a level of one or more glycan structures, wherein the means generates spectral data; a means for generating one or more regression models based on the spectral data; and a means for measuring a level of one or more glycan structures in a glycosylated molecule-producing cell line.

[0029] In one embodiment, the glycosylated molecule-producing cell line is a mammalian cell line. In one embodiment, the mammalian cell line is a non-human cell line.

[0030] In one embodiment, the culturing includes batch, fed-batch, perfusion, of a combination thereof.

[0031] In one embodiment, culturing includes a volume of about 2,000 liters or more. In one embodiment, culturing includes a volume of about 5,000 liters or more. In one embodiment, the bioreactor volume is about 10,000 liters or more. In one embodiment, the bioreactor volume is about 15,000 liters or more. In one embodiment, culturing includes a volume of about 10,000
5 liters to about 25,000 liters. In one embodiment, culturing includes a volume of about 15,000 liters.

[0032] In one embodiment, measuring is performed in-line, at-line, on-line, off-line, or a combination thereof. In one embodiment, measuring is performed in-line. In one embodiment, measuring is performed on-line. In one embodiment, measuring is performed at-line. In one
10 embodiment, measuring is performed off-line.

[0033] In one embodiment, measuring occurs more than once daily. In one embodiment, measuring occurs about every 5 to 60 minutes. In one embodiment, measuring occurs about every 10 to 30 minutes. In one embodiment, measuring occurs about every 10 to 20 minutes. In one embodiment, measuring occurs about every 12.5 minutes.

[0034] In one embodiment, the one or more glycan structures are selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.
15

[0035] In one embodiment, the sytem further includes a means for selectively modifying one or more operating parameters to enhance production of a desired one or more glycan structures. In one embodiment, the one or more operating parameters include a pH level, a nutrient level, a
20 culture media concentration, a frequency interval for culture media addition, or a combination thereof. In one embodiment, the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. In one embodiment, the concentration of glucose is automatically modified based on the spectral data.

[0036] In one embodiment, the one or more glycosylated molecules include a monoclonal antibody (mAb). In one embodiment, the one or more glycosylated molecules include a non-mAb.
25

[0037] In one embodiment, the spectral data comprises Raman spectra.

[0038] In one embodiment, the one or more regression models comprise a partial least squares (PLS) model.

[0039] In one embodiment, the system further includes a means for isolating the one or more glycosylated molecules.

5 [0040] In one aspect, provided herein is a system for producing one or more glycosylated molecules, and the system includes: a bioreactor comprising a glycosylated molecule-producing cell line; a process analytical technology (PAT) tool that measures one or more glycan structures and generates spectral data; and a processor that correlates levels of one or more glycan structures with the spectral data using one or more regression models.

10 [0041] In one embodiment, the bioreactor is about 2,000 liters or more. In one embodiment, the bioreactor is about 5,000 liters or more. In one embodiment, the bioreactor volume is about 10,000 liters or more. In one embodiment, the bioreactor volume is about 15,000 liters or more. In one embodiment, the bioreactor is about 10,000 liters to about 25,000 liters. In one embodiment, the bioreactor is about 15,000 liters.

15 [0042] In one embodiment, the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[0043] In one embodiment, the glycosylated molecule-producing cell line is a mammalian cell line. In one embodiment, the mammalian cell line is a non-human cell line.

20 [0044] In one embodiment, the PAT tool utilizes or otherwise comprises Raman spectroscopy.

[0045] In one embodiment, the one or more regression models comprise a partial least squares (PLS) model.

25 [0046] Other aspects, features and advantages of the invention will be apparent from the following disclosure, including the detailed description of the invention and its preferred embodiments *and the appended claims*.

4. BRIEF DESCRIPTION OF THE DRAWINGS

[0047] The foregoing summary, as well as the following detailed description of preferred embodiments of the present application, will be better understood when read in conjunction with

the appended drawings. It should be understood, however, that the application is not limited to the precise embodiments shown in the drawings.

[0048] FIG. 1A and FIG. 1B illustrate the Maillard reaction for the glycation of mAbs (FIG. 1A), and common N-Glycan structures associated with N-linked glycosylation of therapeutic proteins such as mAbs (FIG. 1B).

[0049] FIG. 2 illustrates a process flow for the development of and testing of Raman based PLS models.

[0050] FIG. 3A – FIG. 3G illustrate Raman PLS model predictions (fluctuating blue line) vs. offline measurements (steady red line) for mono-glycation (FIG. 3A), non-glycation (FIG. 3B), G0F-GlcNac (FIG. 3C), G0 (FIG. 3D), G0F (FIG. 3E), G1F (FIG. 3F), and G2F (FIG. 3G) in reduced scale (5L) batch (CTSS).

[0051] FIG. 4A – FIG. 4G illustrate Raman PLS model predictions (fluctuating blue line) vs. offline measurements (steady red line) for mono-glycation (FIG. 4A), non-glycation (FIG. 4B), G0F-GlcNac (FIG. 4C), G0 (FIG. 4D), G0F (FIG. 4E), G1F (FIG. 4F), and G2F (FIG. 4G) in manufacturing scale (2,000L) batch (CTSS) without adding or removing additional data from the models.

[0052] FIG. 5A – FIG. 5G illustrate Raman PLS model predictions (fluctuating blue line) vs. offline measurements (steady red line) for mono-glycation (FIG. 5A), non-glycation (FIG. 5B), G0F-GlcNac (FIG. 5C), G0 (FIG. 5D), G0F (FIG. 5E), G1F (FIG. 5F), and G2F (FIG. 5G) in manufacturing scale (2000L) batch (CTSS) models using models supplemented with process data from a 2000L scale batch run.

[0053] FIG. 6 illustrates an exemplary system for process flow diagram 600 illustrating an exemplary system used to provide the real time monitoring of glycation and/or glycosylation profiles of therapeutic proteins in manufacturing scale bioreactors.

[0054] FIG. 7 illustrates a process flow diagram 700 depicting the determination of a glycan structure on a glycosylated molecule.

[0055] FIG. 8 illustrates a process flow diagram 800 depicting the production of a glycosylated molecule having a desired glycan structure.

[0056] FIG. 9 illustrates a process flow diagram 900 depicting the determination of glycation on a molecule.

[0057] FIG. 10 illustrates a process flow diagram 1000 depicting production of a molecule with a desired level of glycation.

5 [0058] FIG. 11A – FIG. 11F illustrate an overlay of trends for variable importance of projections (VIP) scores of PLS models for (FIG. 11A) monoglycation/non-glycation, (FIG. 11B) G0F-GlcNac, (FIG. 11C) G0, (FIG. 11D) G0F, (FIG. 11E) G1F, and (FIG. 11F) G2F (in flow 1 and flow 3).

[0059] FIG. 12A – FIG. 12B illustrate plot of Raman glucose values vs offline glucose
10 values for CSS of Glucose model for (FIG. 12A) Cell Line 1 and (FIG. 12B) Cell Line 2.

[0060] FIG. 13A – FIG. 13C illustrate Raman PLS model predictions (fluctuating blue line) vs. offline measurements (steady red line) for Cell Line 1, (FIG. 13A) Batch A: Bolus Fed, (FIG. 13B) Batch B: automated feedback control to 1 g/l, (FIG. 13C) Batch C: automated feedback control to 1 g/l.

15 [0061] FIG. 14A – FIG. 14C illustrate Raman PLS model predictions (fluctuating blue line) vs. offline measurements (steady red line) for Cell Line 2, (FIG. 14A) Batch D: Bolus Fed, (FIG. 14B) Batch E: automated feedback control to 2 g/l, (FIG. 14C) Batch F: automated feedback control to 2 g/l.

[0062] FIG. 15A – FIG. 15H illustrate comparison of process trends for Cell Line 1 bolus
20 fed and automated feedback control batches (FIG. 15A), VCD (FIG. 15B), Viability % (FIG. 15C), LDH (FIG. 15D), pH (FIG. 15E), O₂ (FIG. 15F), Glucose Feed volume (FIG. 15G), Titer and (FIG. 15H) Glycation.

[0063] FIG. 16A – FIG. 16H illustrate comparison of process trends for Cell Line 2 bolus
25 fed and automated feedback control batches (FIG. 16A), VCD (FIG. 16B), Viability % (FIG. 16C), LDH (FIG. 16D), pH (FIG. 16E), O₂ (FIG. 16F), Glucose Feed volume (FIG. 16G), Titer and (FIG. 16H) Glycation.

5. DETAILED DESCRIPTION

[0064] Current methods for measuring glycation and glycosylation profiles include Boronate affinity chromatography, capillary isoelectric focusing colorimetric assays and liquid

chromatography mass spectrometry (LC/MS). These methods, while capable of providing precise and accurate results are both time and resource consuming. Moreover, sampling often comes at the expense of product being removed from the bioreactor and a greater risk of contamination. Another drawback to the current methods is that they involve product quality testing once a batch has been completed.

[0065] Accordingly, there is an unmet need for methods and systems that can accurately monitor glycation and/or glycosylation throughout production of a therapeutic protein, which can enable a user with the ability to control glycation and/or glycosylation during the production process in real time.

[0066] The present disclosure is directed, in part, to methods and systems to control and monitor glycation and/or glycosylation of molecules (such as a therapeutic protein disclosed in Section 5.1) during the production process. For example, methods and systems are described herein that involve the partnering of a process analytical technology (PAT) tool and chemometric modelling to develop predictive models capable of monitoring glycation and glycosylation profiles, including individual glycoforms (microheterogeneity), with a specific focus on manufacturing scale processes. The present disclosure is also directed, in part, to the discovery that by including manufacturing scale data in the chemometric modelling, the predictive power and robustness of the model can be improved. Such methods and systems enable potential quality issues in the therapeutic protein production process to be identified before they impact the batch, and can also help reduce process variability, reduce supply costs due to yield improvements, carry out real time release of the product, reduce lead times and positively impact the technology transfer timelines of products due to a reduction in analytical method transfer and validation processes.

[0067] In some aspects, provided herein is a method for determining a glycan structure on a glycosylated molecule using PAT tool, such as Raman spectroscopy, and generating a regression model that can be used by a computing device to determine a level of one or more glycan structures on the therapeutic protein within a bioreactor. In other aspects, provided herein is a method for producing a glycosylated molecule having a desired glycan structure using a PAT tool that generates spectral data, and determining a level of one or more glycan structures on the therapeutic protein within a bioreactor by using one or more regression models. One or more

operating parameters can then either be maintained or selectively modified based on the levels of the desired and/or undesired glycan structure. The present disclosure also provides a system for producing a one or more glycosylated molecules. The system can include a bioreactor comprising a cell line capable of producing the glycosylated molecule, a PAT tool that measures one or more glycan structures and generates spectral data, and a processor that correlates levels of one or more glycan structures with the spectral data using one or more regression models.

[0068] In some aspects, provided herein is a method for determining glycation on a molecule using a process analytical technology (PAT) tool, such as Raman spectroscopy, and generating a regression model that can be used by a computing device to determine glycation levels on a molecule within a bioreactor. In other aspects, provided herein is a method of producing a molecule with a desired level of glycation by measuring glycation on the molecule using a PAT tool that generates spectral data, and determining levels of glycation on the molecule by using one or more regression models. One or more operating parameters can then be either maintained or selectively modified based on the levels of glycation. The present disclosure also provides a system for producing a non-glycated molecule. The system can include a bioreactor comprising a cell line capable of producing the non-glycated molecule, a PAT tool that measures glycation and generates spectral data, and a processor that correlates levels of glycation with the spectral data using one or more regression models.

[0069] Various publications, articles and patents are cited or described in the background and throughout the specification; each of these references is herein incorporated by reference in its entirety. Discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is for the purpose of providing context for the invention. Such discussion is not an admission that any or all of these matters form part of the prior art with respect to any inventions disclosed or claimed.

[0070] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention pertains. Otherwise, certain terms used herein have the meanings as set forth in the specification. All patents, published patent applications and publications cited herein are incorporated by reference as if set forth fully herein.

[0071] It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise.

[0072] Throughout this specification and the claims that 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 integer or step. When used herein the term “comprising” can be substituted with the term “containing” or “including” or sometimes when used herein with the term “having”.

[0073] When used herein “consisting of” excludes any element, step, or ingredient not specified in the claim element. When used herein, “consisting essentially of” does not exclude materials or steps that do not materially affect the basic and novel characteristics of the claim. Any of the aforementioned terms of “comprising”, “containing”, “including”, and “having”, whenever used herein in the context of an aspect or embodiment of the application can be replaced with the term “consisting of” or “consisting essentially of” to vary scopes of the disclosure.

[0074] As used herein, the conjunctive term “and/or” between multiple recited elements is understood as encompassing both individual and combined options. For instance, where two elements are conjoined by “and/or,” a first option refers to the applicability of the first element without the second. A second option refers to the applicability of the second element without the first. A third option refers to the applicability of the first and second elements together. Any one of these options is understood to fall within the meaning, and therefore satisfy the requirement of the term “and/or” as used herein. Concurrent applicability of more than one of the options is also understood to fall within the meaning, and therefore satisfy the requirement of the term “and/or.”

[0075] As used herein, the term “molecule” generally refers to a protein, or fragment thereof. An exemplary molecule is a therapeutic protein (*e.g.*, a recombinant protein or mAb) or fragment thereof.

[0076] As used herein, the term “bioreactor” generally refers to a device that supports a biologically active process, such as the culture of cells. Exemplary bioreactors include stainless steel stirred tank bioreactors, air-lift reactors, and disposable bioreactors.

[0077] As used herein, the term “spectral data” generally refers to the analytical output obtained using spectroscopy.

[0078] As used here, the term “threshold” generally refers to an amount that defines at least one limit, such as a lower limit or an upper limit, of a particular range or level. The threshold can
5 be determined empirically, or it can be a pre-defined threshold that is set *a priori*.

[0079] As used herein, the term “selectively modifying” when used in reference to an operating parameter, generally refers to the purposeful adjustment of one or more conditions to facilitate optimal production of the desired product and/or decrease production of an undesired product.

10 [0080] As used herein, the term “automatically modified” generally means that an operating parameter is adjusted without the user needing to manually adjust the operating parameter.

[0081] As used herein, the term “on-site” generally means that the measuring or determining is performed at the same facility where production occurs.

[0082] As used herein, the term “off-site” generally means that the measuring or determining
15 is performed at a facility different from where production occurs.

[0083] As used herein, the term “off-line” generally means that the measuring or determining is performed after the production process has complete, and collection of the sample is manual.

[0084] As used herein, the term “at-line” generally means the measuring or determining is performed within the production area, and collection of the sample is manual.

20 [0085] As used herein, the term “on-line” generally means that the measuring or determining is performed within the production area, and collection of the sample is automated.

[0086] As used herein, the term “in-line” generally means that the measuring or determining is performed in real time within the production area by a probe placed in the bioreactor and collection of a sample is not required.

25 [0087] The practice of the embodiments provided herein will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, and immunology, which are within the skill of those working in the art. Such techniques are explained fully in the literature. Examples of particularly suitable texts for consultation include the following:

Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Third Ed., Cold Spring Harbor Laboratory, New York (2001); Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore, MD (1999); Glover, ed., DNA Cloning, Volumes I and II (1985); Freshney, ed., Animal Cell Culture: Immobilized Cells and Enzymes (IRL Press, 1986); Källen
5 *et al.*, Plant Molecular Biology – A Laboratory Manual (Ed. by Melody S. Clark; Springer-Verlag, 1997); Immunochemical Methods in Cell and Molecular Biology (Academic Press, London); Scopes, Protein Purification: Principles and Practice (Springer Verlag, N.Y., 2d ed. 1987); National Research Council (US) Committee on Methods of Producing Monoclonal Antibodies. Monoclonal Antibody Production. Washington (DC): National Academies Press
10 (US); 1999; and Clausen H, *et al.* Glycosylation Engineering. 2017. In: Varki A, Cummings RD, Esko JD, *et al.*, editors. Essentials of Glycobiology. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017; and National Research Council (US) Committee on Revealing Chemistry through Advanced Chemical Imaging. Visualizing Chemistry: The Progress and Promise of Advanced Chemical Imaging. Washington (DC): National Academies
15 Press (US); 2006. 3, Imaging Techniques: State of the Art and Future Potential.

[0088] In an attempt to help the reader of the application, the description has been separated in various paragraphs or sections, or, is directed to various embodiments of the application. These separations should not be considered as disconnecting the substance of a paragraph or section or embodiments from the substance of another paragraph or section or embodiments. To
20 the contrary, one skilled in the art will understand that the description has broad application and encompasses all the combinations of the various sections, paragraphs and sentences that can be contemplated. The discussion of any embodiment is meant only to be exemplary and is not intended to suggest that the scope of the disclosure, including the claims, is limited to these examples. The application contemplates use of any of the applicable components in any
25 combination, whether or not a particular combination is expressly described.

5.1. Therapeutic proteins

[0089] The present disclosure is related, in part, to determining and/or measuring glycation and/or glycosylation of molecule. In certain embodiments, the molecule is a therapeutic protein. Non-limiting examples of therapeutic proteins include, for example, antibody-based drugs (*e.g.*,
30 polyclonal antibodies, or monoclonal antibodies (mAb)), Fc fusion proteins, anticoagulants,

blood factors, bone morphogenetic proteins, engineered protein scaffolds, enzymes, growth factors, hormones, interferons, interleukins, recombinant proteins, and thrombolytics.

Accordingly, in some embodiments, the molecule is a mAb. In some embodiments, the molecule is a recombinant protein. In some embodiments, the molecule is an Fc fusion protein. In some
5 embodiments, the molecule is an anticoagulant. In some embodiments, the molecule is a blood factor. In some embodiments, the molecule is a bone morphogenetic protein. In some
embodiments, the molecule is an engineered protein scaffold. In some embodiments, the molecule is an enzyme. In some embodiments, the molecule is a growth factor. In some
embodiments, the molecule is a hormone. In some embodiments, the molecule is an interferon.
10 In some embodiments, the molecule is an interleukin. In some embodiments, the molecule is a thrombolytic.

5.2. Glycation

[0090] The present disclosure is related, in part, to determining and/or measuring glycation of therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1). Protein glycation
15 is a non-enzymatic glycosylation on protein amine groups that generally occurs on the alpha amine terminal and epsilon amine group on the lysine side chain. Glycation involves a process
by which reducing sugars, such as glucose, fructose and galactose, covalently binds to a protein in a non-enzymatic reaction. The reaction between the amino acid and reducing sugar was first
described by Maillard in 1912. The susceptible amine group reversibly condenses with an
20 aldehyde group of the reducing sugar, to form an unstable Schiff base intermediate, which can
undergo a spontaneous multistep Amadori rearrangement to form a more stable, covalently
bonded ketoamine. This reaction results in irreversible products causing biophysical and
structural changes in proteins.

[0091] Glycation can potentially affect bioactivity and/or molecular stability of therapeutic
25 proteins. For example, in the case of mAbs, which represent an exemplary therapeutic protein of
the present disclosure, glycation can block the biologically functional site and/or cause
degradation of the mAbs. The degradation may further lead to aggregation of the mAb.
Consequently, glycation of therapeutic proteins represents a potential critical quality attribute
(CQA) to the production process.

5.3. Glycosylation

[0092] The present disclosure is also related, in part, to determining, and/or measuring glycosylation, including specific glycan structures, of therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1).

5 [0093] Glycosylation is a complex post translation modification that involves attachment of glycans to specific sites (*e.g.*, N-linked and/or O-linked glycosylation) on the therapeutic protein. N-linked glycosylation of mAbs, for example, generally involves attachment of glycans at the Asn-X-Ser/Thr sequence of the Fc portion of mAb heavy chains, where X can be any amino acid except for proline (Ghaderi *et al.*, 2012). For example, therapeutic proteins can include IgG1
10 molecules, which contain a single N-linked glycan at Asn297 in each of the two heavy chains. N-glycosylation can also occur in the variable region of each heavy chain (*e.g.*, cetuximab). O-linked protein glycosylation generally involves a linkage between the monosaccharide N-Acetylgalactosamine and the amino acids serine or threonine.

[0094] Depending on the arrangement of carbohydrate moieties covalently linked to specific
15 regions of the therapeutic protein, classes of these carbohydrates (called “glycans”) can be derived. For example, N-glycans include G0, G0F, G0F-GlcNac, G1F and G2F, which differ slightly in structure (see FIG 1B). Variations in the structure of these glycans result in subtle changes in structure that can have a significant influence on protein activity, conformation, safety, and efficacy. As a result, glycosylation of therapeutic proteins has a pivotal role in their
20 safety and efficacy, including immunogenicity, and appropriate glycosylation is one of the critical quality attributes (CQAs) that must be demonstrated to ensure the safety and potency of commercial therapeutic proteins, such as mAbs, before regulatory approval.

5.4. Process Analytical Technology (PAT)

[0095] Current methods for measuring glycation and glycosylation profiles include Boronate
25 affinity chromatography, capillary isoelectric focusing colorimetric assays and liquid chromatography mass spectrometry (LC/MS). These methods, while capable of providing precise and accurate results are both time and resource consuming. Moreover, sampling using these methods often comes at the expense of product being removed from the bioreactor and a greater risk of contamination. Another drawback to the current methods is that they involve
30 product quality testing once a batch has been completed.

[0096] As provided herein, the present disclosure relates, in part, to the measurement of glycation and/or glycosylation using Process Analytical Technology (PAT). PAT is advantageous because it can be used to control and monitor the production process before the batch has been completed. For example, PAT can allow samples to be measured in real time, such as by measuring in-line, or otherwise measured during the production process, such as by measuring at-line or on-line. Real time monitoring, for example, can increase process control, as it provides the opportunity to identify potential quality issues before they impact the batch, reduce process variability, reduce supply costs due to yield improvements, carry out real time release of the product, reduce lead times and positively impact the technology transfer timelines of products due to a reduction in analytical method transfer and validation processes. In contrast, complex and time-consuming methods that are reliant on the frequency at which off-line samples are measured only give a limited understanding of these CQAs, such as glycation and glycosylation, and are often only assessed after the bioreactor process has been completed. Therefore, by applying PAT during the production process, the quality of the product can be more carefully monitored and the final product produced will be a higher quality, as compared to production that relies on end-of-line testing to filter out the products that are outside their specifications.

5.4.1. PAT tools

[0097] As provided herein, detection and/or measuring glycation and/or glycosylation can be carried out using PAT tools that involve spectroscopic techniques such as, for example, fluorescence spectroscopy, diffuse reflectance spectroscopy, infrared spectroscopy (*e.g.*, near-infrared, or mid-infrared), terahertz spectroscopy, transmission and absorbance spectroscopy, Raman spectroscopy, including Surface Enhanced Raman Spectroscopy (SERS), Spatially Offset Raman spectroscopy (SORS), transmission Raman spectroscopy, and/or resonance Raman spectroscopy.

[0098] By way of example, Raman spectroscopy is a PAT tool capable of providing data that can be partnered with chemometric modelling (such as the chemometric modelling described in Section 5.4.2). It provides clear, sharp spectra, and can be optionally recorded within the bioreactor (*e.g.*, *in situ*) using an in-situ probe during the production processes. Raman

spectroscopy is a vibrational spectroscopic technique that uses laser technology to provide a chemical fingerprint of a substance.

[0099] Moreover, it has been used as a PAT tool to provide non-destructive real time measurement of a number of biotherapeutic process variables including metabolites, growth profiles, product levels, product quality attributes, nutrient feeds and more recently, culture pH. It should be understood, however, that other spectroscopic techniques can also be used with the methods and systems provided herein, and that Raman spectroscopy is merely an exemplary PAT tool.

[00100] Spectra obtained from any of the PAT tools described herein can be collected with a probe that is within the bioreactor (*i.e., in situ*), or by collecting a sample from the bioreactor and measuring the sample outside of the bioreactor. Spectra can also be coupled with multivariate analysis (MVA) to allow for monitoring of other operating parameters in addition to glycation and/or glycosylation such as, for example, metabolites, and/or cell concentration. For example, data on one or more operating parameters (*e.g.*, pH, temperature, pressure, dissolved oxygen, optical density, oxygen uptake rate, *etc.*) can be combined with supporting information (raw material analysis, timing and duration of feeds, manual cell counts, metabolite levels, *etc.*) to generate large amounts of high dimensional data on the production process that can be handled by chemometric modeling methods such as the chemometric modeling described in Section 5.4.2). Spectra can also be collected pre and post glucose feeding.

[00101] Therefore, by combining spectroscopic analysis with chemometric modeling methods (such as the chemometric modeling described in Section 5.4.2), the real time monitoring of glycation and/or glycosylation, and optionally additional information on the production process, can be achieved. Therefore, spectroscopy, such as Raman spectroscopy, provides the ability to monitor bioprocesses during the production process rather than after the run is completed. As a result, spectroscopy can be used as a PAT tool for measuring glycation and/or glycosylation, with the option to implement the measurement results as a feedback loop that controls one or more operating parameters such as, for example, nutrient feeds, thereby leading to the desired amount of glycation and/or glycosylation, and optionally the specific glycan structure, on the therapeutic protein.

5.4.2. Chemometric modeling

[00102] The spectroscopic methods provided herein can generate large amounts of high dimensional data. Generally, the data is handled by chemometric modeling methods using known techniques such as, for example, partial least square (PLS), classic least squares (CLS) or principle component analysis (PCA). The model is then built with the captured absorption spectra and reference levels obtained off-line, such as by liquid chromatography–mass spectrometry (LC–MS). The spectra can be subject to preprocessing methodologies, such as first and second or derivatives, extended multiplicative scattering correction, mean centering, and auto scaling, to name a few. The preprocessing methodologies can be used to help mitigate interferences such as cloudiness, or optical transmissibility, of the fluid, instrument drift, and contaminate build up on the lenses in contact with the fluid. The preprocessing methodologies also act as noise filters to enable models to focus on the real compositional changes in the fluid that may affect the resultant vapor pressure of the liquid. After that, the chemometric model is implemented to the PAT tool analyzer as the calibration curve to predict the glycation and/or glycosylation in real time.

[00103] Accordingly, in some embodiments, the one or more regression models is selected from the group consisting of partial least square (PLS), classic least squares (CLS) and principle component analysis (PCA). In some embodiments, the one or more regression models includes a partial least squares (PLS) model. In some embodiments, the one or more regression models includes a classic least squares (CLS) model. In some embodiments, the one or more regression models includes a principle component analysis (PCA).

5.4.3. Measurement parameters.

[00104] As provided herein, glycation and/or glycosylation can be measured using a PAT tool, such as Raman spectroscopy, during the production process (*e.g.*, at-line, on-line, and/or in-line), or after the production process is completed (*e.g.*, off-line). Accordingly, in some embodiments, measuring is performed in-line, at-line, on-line, off-line, or a combination thereof. In some embodiments, the measuring is performed in-line. In some embodiments, the measuring is performed at-line. In some embodiments, the measuring is performed on-line. In some embodiments, the measuring is performed off-line.

[00105] Glycation and/or glycosylation can also be optionally measured on-site or off-site. In some embodiments, the measuring is performed on-site. In some embodiments, it is performed off-site.

[00106] The PAT tools disclosed herein that provide spectra, are capable of providing frequent measurement during the production process. For example, by using Raman spectroscopy, there is the potential to deliver a predicted value for a range of process variables normally confined to a single daily offline measurement as frequently as once every 15 minutes, when a single probe is in use, generating up to 360 times the amount of information versus a single daily offline measurement for a 16 day process.

[00107] Accordingly, in some embodiments measurement can occur more than once daily. If more frequent measurement is desired, measuring can occur about every 5 to 60 minutes, about every 10 to 30 minutes, about every 10 to 20 minutes, or about every 12.5 minutes. Generally, more frequent the measurements will be able to allow for more precise predictions of the glycation and/or glycosylation levels. In turn, this will allow the user ability to better control the molecule's glycation and/or glycosylation levels.

5.5. Production

[00108] As provided herein, the present disclosure relates, in part, to measuring glycation and/or glycosylation on therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1) that are produced by culturing a biologically active organism (such as by using the cell culture methods described in Section 5.5.2) in bioreactors (such as the bioreactors describe in Section 5.5.1).

5.5.1. *Bioreactors*

[00109] As provided herein, the present disclosure relates, in part, to measuring glycation and/or glycosylation on therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1) that are grown in bioreactors. Various types of bioreactors can be used for production of the therapeutic protein. For example, the bioreactor can be a stainless steel stirred tank bioreactor (STR), an air-lift reactor, a disposable bioreactor, or a combination thereof (e.g., a disposable bioreactor combined with the STR).

[00110] The bioreactor can have any suitable volume that allows for the cultivation and propagation of biological cells capable of producing therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1). For example, the volume of the bioreactor can be about 0.5 liters (L) to about 25,000 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 250 L. In some embodiments, the volume of the bioreactor can be about 0.5 liters (L) to about 250 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 50 L. In some embodiments, the volume of the bioreactor can be about 1 L to about 50 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 25 L. In some embodiments, the volume of the bioreactor can be about 1 L to about 25 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 10 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 5 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 1 L. In some embodiments, the volume of the bioreactor can be about 1 L. In some embodiments, the volume of the bioreactor can be about 2 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 5 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 10 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 25 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 50 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 100 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 250 L. In some embodiments, the volume of the bioreactor can equal to or above 1,000 L. In some embodiments, the volume of the bioreactor can be about 1,000 L to about 25,000 L. In some embodiments, the volume of the bioreactor can be about 10,000 L to about 25,000 L. In some embodiments, the volume of the bioreactor can be about 1,000 L. In some embodiments, the volume of the bioreactor can be about 2,000 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 5,000 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 10,000 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 15,000 L. In some embodiments, the volume of the bioreactor can be less than or equal to about 25,000 L.

[00111] As provided herein, in certain aspects of the present disclosure, chemometric models can be first generated using data obtained from reduced scale bioreactors (*e.g.*, less than 250 L), and the robustness and predictive power of the chemometric models can be increased by refining

the model using data obtained from the manufacturing scale bioreactor (*e.g.*, greater than or equal to 2,000 L, and preferably between 10,000 L to 25,000 L) for use in predicting glycation and/or glycosylation at manufacturing scale. Accordingly, in some embodiments, the methods and systems described herein involve two or more bioreactors having different volumes.

5 **[00112]** The addition of data obtained from the manufacturing scale bioreactor improves the model's ability to predict. This may be due to the changes in the bioreactor culture that occur in the manufacturing scale bioreactor, but not in the reduced scale process. For example, culture mixing time, CO₂ removal and oxygen transfer rates can differ with the manufacturing scale. By way of example, a reduced oxygen transfer in the manufacturing scale process has the potential
10 to influence the glycation and, in particular, the glycosylation profiles of the mAb products in a way that would not be seen at the reduced scale. A reduced oxygen transfer leads to potential presence of 'dead zones' areas in the bioreactor where no oxygen is present for brief periods (Ast *et al.*, 2019). This leads to an increased oxidative stress on the host cells in the bioreactor. Oxidative stress has been reported to have an effect on the glycosylation of mAbs as this stress
15 reduces acetyl-CoA formation which in turn leads to a decreased N-acetylglucosamine (GlcNac) (Lewis *et al.*, 2016), a key amide which forms part of the backbone structure of glycosylation targets described here, apart from G0F-GlcNac. Therefore, by incorporating data obtained from the manufacturing scale process, these variations can be accounted for in the model, and help improve the model's predictability.

20 **[00113]** Thus, in some embodiments, the chemometric model involves data generated from one or more first bioreactors, as well as data generated from one or more second bioreactors. In some embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 250 L. In some embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 100 L. In some
25 embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 50 L. In some embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 25 L. In some embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 10 L. In some embodiments, each of the one or more first bioreactors has a maximum threshold
30 volume of less than or equal to about 5 L. In some embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 2 L. In some

embodiments, each of the one or more first bioreactors has a maximum threshold volume of less than or equal to about 1 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 250 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 100 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 50 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 25 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 10 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 5 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 2 L. In some embodiments, each of the one or more first bioreactors has a volume of less than or equal to about 1 L. In some embodiments, the volume of each of the one or more first bioreactors can be about 0.5 liters (L) to about 250 L. In some embodiments, the volume of each of the bioreactors can be about 1 L to about 50 L. In some embodiments, the volume of each of the one or more first bioreactors can be about 1 L to about 25 L. In some embodiments, the volume of each of the one or more first bioreactors can be about 1 L to about 10 L. In some embodiments, the volume of each of the one or more first bioreactors can be about 1 L to about 5 L.

[00114] In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 1,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 2,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 5,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 10,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 15,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 20,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of greater than or equal to about 25,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of about 2,000 L to about 25,000 L. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume of about 5,000 L to about 25,000 L. In some embodiments, each of

the one or more second bioreactors has a minimum threshold volume of about 10,000 L to about 25,000 L. In a preferred embodiment, each of the one or more bioreactors has a minimum threshold volume equal to the volume of the bioreactor used for manufacturing the therapeutic protein. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 1,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 2,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 5,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 10,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 15,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 20,000 L. In some embodiments, each of the one or more second bioreactors has a volume of greater than or equal to about 25,000 L. In some embodiments, each of the one or more second bioreactors has a volume of about 1,000 L to about 25,000 L. In some embodiments, each of the one or more second bioreactors has a volume of about 2,000 L to about 25,000 L. In some embodiments, each of the second bioreactors has a volume of about 5,000 L to about 25,000 L. In some embodiments, each of the second bioreactors has a volume of about 10,000 L to about 25,000 L. In some embodiments, each of the one or more second bioreactors has a volume of about 15,000 L to about 25,000 L. In a preferred embodiment, each of the one or more bioreactors has a volume equal to the volume of the bioreactor used for manufacturing the therapeutic protein.

[00115] In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least five times (5X) greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 10X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 25X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 50X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 100X greater than the maximum threshold volume of each of the

one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 250X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 500X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a minimum threshold volume that is at least 1,000X greater than the maximum threshold volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least five times (5X) greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 10X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 25X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 50X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 100X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 250X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 500X greater than the volume of each of the one or more first bioreactors. In some embodiments, each of the one or more second bioreactors has a volume that is at least 1,000X greater than the volume of each of the one or more first bioreactors.

5.5.2. *Cell culture methods*

[00116] The production of the therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1) of the present disclosure can be performed with any suitable biologically active host-cell type known in the art to be capable of producing therapeutic proteins. For example, mammalian cells and non-mammalian cells can be used as platforms for the production of therapeutic proteins. Non-limiting examples of mammalian host cell lines for production of therapeutic proteins include the Chinese hamster ovary (CHO), mouse myeloma derived NS0 and Sp2/0 cells, human embryonic kidney cells (HEK293), and human embryonic retinoblast-derived PER.C6 cells. Non-limiting examples of non-mammalian hosts include, for example,

Pichia pastoris, *Arabidopsis thaliana*, *Nicotiana benthamiana*, *Aspergillus niger*, and *Escherichia coli*. Accordingly, in some embodiments, the therapeutic protein-producing cell line is a mammalian cell line. In some embodiments, the mammalian cell line is a non-human cell line.

5 **[00117]** Different host systems may express varying glycosylation enzymes and transporters, contributing to the specificity and heterogeneity in glycosylation profiles of the therapeutic protein. Similarly, different host systems may have different effects on the glycation levels of the therapeutic protein. Therefore, it is possible to engineer or specifically select a host-cell line that can generate the desired therapeutic protein with, for example, a specific glycan structure and/or
10 level of glycation.

[00118] Cultivation and propagation of the biologically active cells can be performed using various methods known in the art, such as batch process, fed-batch process, continuous culture, or a combination thereof (*e.g.*, a hybrid between fed batch and perfusion).

[00119] In the batch method, all nutrients are supplied at the beginning of the cultivation,
15 without adding any more in the subsequent bioprocess. During the entire bioprocess, no additional nutrients are added, but control elements such as gases, acids and bases may optionally be added. The bioprocess then lasts until the nutrients are consumed.

[00120] The fed-batch method adds nutrients into the bioreactor as a large bolus at set time points or once they are depleted. Generally, the same medium used in the initial culture is also
20 used for feeding, but in a more concentrated version. The feeding solution composition can be designed to supply the cells based on their metabolic state at different culture phases such as by, for example, analyzing and identifying the spent medium nutrients that are being consumed at greater rates. In addition, the medium used in the fed-batch method can even be modified to accommodate the needs of the cell culture or promote production of the therapeutic protein, such
25 as to promote cell growth or to stimulate production of the therapeutic protein. For example, in some embodiments, the medium can be modified to reduce or eliminate glycation of the therapeutic protein.

[00121] In continuous culture, nutrients are continuously added to the bioreactor and, generally, an equivalent amount of converted nutrient solution is simultaneously taken out of the
30 system. The three most common types of continuous culture are chemostat, turbidostat, and

perfusion. The perfusion method circulates medium through a growing culture, allowing simultaneous removal of waste, supply of nutrients, and harvesting of product.

[00122] As provided herein, in some embodiments, the feeding is automated. In certain aspects, the feeding is automated and automation controls one or more operating parameters (such as the operating parameters described in Section 5.5.3).

5.5.3. *Operating parameters*

[00123] Glycation of therapeutic proteins can occur during the production process, where a reducing sugar, such as glucose, is used as an energy source for a biologically active organism, such as a cell culture that produces the therapeutic protein. The level of glycation can be affected by the amount of sugar that is added to the cell culture during the mammalian cell culture process. In addition, other factors, such as pH level, nutrient levels, time (*e.g.*, the frequency interval for culture media addition), temperature, and ionic strength, can affect the kinetics and extent of glycation. Moreover, the specific types of sugars used in the cell culture media, such as hexose sugars, for example, and the specific reactivity of the accessible amino groups can affect the protein glycation.

[00124] Many process parameters can shape the glycosylation of therapeutic proteins such as, for example, dissolved oxygen (DO) levels, culture temperatures of bioreactor, host-cell type, and nutrient or supplement availability.

[00125] Accordingly, in some embodiments, one or more operating parameters of the bioreactor is maintained or selectively modified based on the determined glycation and/or glycosylation levels to produce a therapeutic protein with an acceptable amount of glycation and/or glycosylation. In some embodiments, the operating parameter includes a temperature, pH level, a dissolved oxygen (DO) level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof.

[00126] In some embodiments, the operating parameter includes a temperature. For example, if the therapeutic protein has a desirable amount of glycation and/or glycosylation, the temperature of the bioreactor can be maintained at or around the temperature in the bioreactor at the time of measuring the glycation and/or glycosylation level. Alternatively, if the therapeutic protein has an undesirable amount of glycation and/or glycosylation, the temperature of the

bioreactor can be adjusted for a portion or the entirety of the production process to achieve the desired amount of glycation and/or glycosylation. In some embodiments, the temperature is a physiological temperature. In some embodiments, the temperature is maintained or adjusted to a temperature of about 25°C to 42°C. In some embodiments, the temperature is maintained or
5 adjusted to a temperature of about 35°C to 39°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 35.5°C to 37.5°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 35°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 35.5°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 36°C. In some embodiments,
10 the temperature is maintained or adjusted to a temperature of about 36.5°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 37°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 37.5°C. In certain embodiments, the temperature in the bioreactor is a non-physiological temperature. In some embodiments, the temperature is maintained or adjusted to a temperature of less than 25°C.
15 In some embodiments, temperature is maintained or adjusted to a temperature of about 4°C to about 25°C. In some embodiments, temperature is maintained or adjusted to a temperature of about 4°C to about 10°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 4°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 5°C. In some embodiments, the temperature is maintained or adjusted to a
20 temperature of about 6°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 7°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 8°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 9°C. In some embodiments, the temperature is maintained or adjusted to a temperature of about 10°C.

25 **[00127]** In some embodiments, the operating parameter includes a pH level that is either maintained or modified, depending on the measured level of glycation and/or glycosylation on the therapeutic protein. For example, if the therapeutic protein has a desirable amount of glycation and/or glycosylation, the pH in the bioreactor can be maintained at or around the pH level in the bioreactor at the time of measuring the glycation and/or glycosylation level. High pH
30 (≥ 7.0) can be beneficial for initial cell growth phase. However, high pH, high lactate and high osmolality cascade often causes delayed cell growth and accelerated cell death. Accordingly, the

pH level may also need to be adjusted during the production process depending on what phase of growth the biologically active cells are in. For example, pH can be maintained or adjusted using addition(s) of carbon dioxide and/or sodium carbonate.

[00128] In some embodiments, the pH is a physiological pH. In some embodiments, the pH is between pH 4.0 to pH 9.0. In some embodiments, the pH is between pH 5.0 to pH 8.0. In some embodiments, the pH is between pH 6.0 to pH 7.0. In some embodiments, the pH is between pH 4.0 to pH 6.0. In some embodiments, the pH is between pH 5.0 to pH 7. In some embodiments, the pH is between pH 6.0 to pH 8.0. In some embodiments, the pH is maintained or adjusted to about pH 4.0. In some embodiments, the pH is maintained or adjusted to about pH 4.5. In some embodiments, the pH is maintained or adjusted to about pH 5.0. In some embodiments, the pH is maintained or adjusted to about pH 5.5. In some embodiments, the pH is maintained or adjusted to about pH 6.0. In some embodiments, the pH is maintained or adjusted to about pH 6.5. In some embodiments, the pH is maintained or adjusted to about pH 6.6. In some embodiments, the pH is maintained or adjusted to about pH 6.7. In some embodiments, the pH is maintained or adjusted to about pH 6.8. In some embodiments, the pH is maintained or adjusted to about pH 6.9. In some embodiments, the pH is maintained or adjusted to about pH 7.0. In some embodiments, the pH is maintained or adjusted to about pH 7.1. In some embodiments, the pH is maintained or adjusted to about pH 7.2. In some embodiments, the pH is maintained or adjusted to about pH 7.3. In some embodiments, the pH is maintained or adjusted to about pH 7.4. In some embodiments, the pH is maintained or adjusted to about pH 7.5. In some embodiments, the pH is maintained or adjusted to about pH 7.6. In some embodiments, the pH is maintained or adjusted to about pH 7.7. In some embodiments, the pH is maintained or adjusted to about pH 7.8. In some embodiments, the pH is maintained or adjusted to about pH 7.9. In some embodiments, the pH is maintained or adjusted to about pH 8.0.

[00129] In some embodiments, the operating parameter includes a culture media glucose concentration. For example, if the therapeutic protein has a desirable level of glycation and/or glycosylation, the culture media glucose concentration that is added to the bioreactor, such as by fed-batch culture or perfusion culture, or hybrid fed-batch and perfusion glucose concentration can be maintained at or around the target concentration in use at the time of measuring the glycation and/or glycosylation level.

[00130] In some embodiments, the operating parameter includes the frequency interval for culture media glucose addition. For example, if the therapeutic protein has a desirable level of glycation and/or glycosylation, the frequency interval for culture media glucose addition to the bioreactor, such as in a fed-batch culture or perfusion culture, or hybrid fed-batch and perfusion glucose concentration can be maintained at or around the frequency at the time of measuring the glycation and/or glycosylation level. In some embodiments, the frequency is maintained or adjusted such that the culture media addition is continuous. In some embodiments, the frequency is maintained or adjusted such that the culture media addition is at split intervals, for example, six hours. In some embodiments, the frequency is maintained or adjusted such that the culture media addition is at split intervals, for example, twelve hours. In some embodiments, the frequency is maintained or adjusted such that the culture media addition is daily bolus, for example, twenty-four hours. In some embodiments, the frequency is maintained or adjusted such that the culture media addition is longer than every twenty-four hours.

[00131] In some embodiments, the operating parameter includes a nutrient level. In some embodiments, the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. In some embodiments, the operating parameter includes a concentration of glucose. In some embodiments, the operating parameter includes a concentration of lactate. In some embodiments, the operating parameter includes a concentration of glutamine. In some embodiments, the operating parameter includes a concentration of ammonium ions.

[00132] In certain embodiments, the glucose concentration is maintained or adjusted to about 0.5 g/L to about 40 g/L, depending on the cell density. In certain embodiments, the glucose concentration is maintained or adjusted to about 0.5 g/L to about 30 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 0.5 g/L to about 20 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 0.5 g/L to about 10 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 0.5 g/L to about 5 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 5 g/L to about 40 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 10 g/L to about 40 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 20 g/L to about 40 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 30 g/L to about 40 g/L. In certain embodiments,

the glucose concentration is maintained or adjusted to about 35 g/L to about 40 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 10 g/L to about 30 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 10 g/L to about 20 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 5 g/L to about 10 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 10 g/L to about 15 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 15 g/L to about 20 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 20 g/L to about 25 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 25 g/L to about 30 g/L. In certain embodiments, the glucose concentration is maintained or adjusted to about 30 g/L to about 35 g/L.

[00133] By way of example, a therapeutic protein that has an acceptable level of glycation can have the concentration of a reducing sugar, such as glucose, maintained. Alternatively, as further example, a therapeutic protein that has an undesirable amount of glycation can have a nutrient level, such as the concentration of a reducing sugar, decreased, and/or the frequency of culture media addition can be decreased. It would be within the skillset of a person skilled in the art to understand how much the operating parameters should be adjusted, if at all. Moreover, other alternative operating parameters can be adjusted according to the glycation levels, and it should be understood that the examples described above are intended to be merely exemplary.

[00134] Similarly, a therapeutic protein that has an acceptable level of a specific glycan structure can have one or more operating parameters maintained. Alternatively, a therapeutic protein that has an undesirable level of a specific glycan structure can have one or more operating parameters adjusted to control the specific glycan structure.

[00135] In certain aspects, the one or more operating parameter automatically modified based on the measure spectral data. For example, glycation and/or glycosylation can be measured using a PAT tool, such as a PAT tool described in Section 5.4.1, and the one or more operating parameters can be automatically maintained in the bioreactor if the level of a desired glycan structure is above a pre-defined threshold, or the level of an undesired glycan structure is below a pre-defined threshold. Alternatively, the one or more operating parameters can be automatically

modified in the bioreactor if the level of a desired glycan structures is below a pre-defined threshold, or the level of an undesired glycan structure is above a pre-defined threshold.

[00136] The threshold of an acceptable level of glycation on the therapeutic protein will generally be determined empirically. In some embodiments, the pre-defined threshold of

glycation will be less than 50% of the therapeutic protein in the bioreactor. In some
embodiments, the pre-defined threshold of glycation will be less than 40%. In some
embodiments, the pre-defined threshold of glycation will be less than 30%. In some
embodiments, the pre-defined threshold of glycation will be less than 25%. In some
embodiments, the pre-defined threshold of glycation will be less than 20%. In some
embodiments, the pre-defined threshold of glycation will be less than 15%. In some
embodiments, the pre-defined threshold of glycation will be less than 10%.

[00137] Accordingly, in some embodiments, at least 90% of the therapeutic protein in the
bioreactor will be non-glycated. In some embodiments, at least 85% of the therapeutic protein in
the bioreactor will be non-glycated. In some embodiments, at least 80% of the therapeutic
protein in the bioreactor will be non-glycated. In some embodiments, at least 75% of the
therapeutic protein in the bioreactor will be non-glycated. In some embodiments, at least 70% of
the therapeutic protein in the bioreactor will be non-glycated. In some embodiments, at least 60%
of the therapeutic protein in the bioreactor will be non-glycated. In some embodiments, at least
50% of the therapeutic protein in the bioreactor will be non-glycated.

[00138] The threshold for the level of glycation on the therapeutic protein will also generally
be determined empirically. For example, some level of variability in the batch to batch
glycoform content in a production setting are to be expected, and generally the threshold will be
an acceptable limit that does not produce too much variability in the glycoform content.

Accordingly, in some embodiments the pre-defined threshold of an undesirable glycan structure
will be less than 50% of the therapeutic protein in the bioreactor. In some embodiments, the pre-
defined threshold of an undesirable glycan structure will be less than 40%. In some
embodiments, the pre-defined threshold of an undesirable glycan structure will be less than 30%.
In some embodiments, the pre-defined threshold of an undesirable glycan structure will be less
than 25%. In some embodiments, the pre-defined threshold of an undesirable glycan structure
will be less than 20%. In some embodiments, the pre-defined threshold of an undesirable glycan

structure will be less than 15%. In some embodiments, the pre-defined threshold of an undesirable glycan structure will be less than 10%.

5.5.4. *Purification.*

[00139] Generally, the therapeutic proteins (such as a therapeutic protein disclosed in Section 5.1) are secreted from the cells grown during the production process. After the production process is completed, the therapeutic protein can be separated from the cells and purified using any technique known in the art suitable for protein purification, including purification according to FDA standards. Preferably, the purification will remove process impurities, such as host cell proteins, nucleic acids, and/or lipids.

[00140] The purification process can involve one or more steps. For example, purification of the therapeutic protein can involve primary recovery, purification, and/or polishing steps. Generally, the primary recovery step consists of centrifugation and/or depth filtration in order to remove the cells and cell debris from the culture broth and clarification of the cell culture supernatant that contains the therapeutic protein product. Additional techniques known in the art can be additionally included to improve in the primary recovery process. For example, the use of flocculants such as simple acids, divalent cations, polycationic polymers, caprylic acid and stimulus-responsive polymers can enhance cell culture clarification and reduce the levels of cells, cellular debris, DNA, host cell proteins (HCP) and/or viruses, while preserving the therapeutic protein within the product stream.

[00141] Purification can also involve one or more chromatography techniques (*e.g.*, affinity, ion exchange, hydrophobic interaction), lectin-based purification, boronate-based purification, and/or filtration techniques (*e.g.*, ultrafiltration) that are used as a capture step in separating the product from smaller impurities and/or as a concentration step in reducing the overall volume.

[00142] Optionally, polishing steps can also be performed in order to, for example, remove viruses, aggregated protein, and any other impurities before storing the therapeutic protein. The polishing steps can include, for example, virus filtration, hydrophobic interaction chromatography, and/or filtration steps (*e.g.*, ultrafiltration/diafiltration and/or sterile filtration).

[00143] Accordingly, in some embodiments, the methods and systems provided herein involve purifying the therapeutic protein.

5.6. Methods and systems for determining glycosylation and/or glycation.

[00144] With reference to diagram 600 of FIG. 6, the current subject matter can utilize a process analytical technology (PAT) tool 610 that can have a probe 612 that is either internal or extends into a bioreactor 620 to monitor or otherwise characterize aspects of reactions occurring within the bioreactor 620. The PAT tool 610 comprises one or more data processors and memory into which instructions can be loaded and executed by the data processors. The PAT tool 610 can take various forms and, in some cases, utilizes or otherwise comprises Raman spectroscopy to characterize aspects of the reactions occurring within the bioreactor 620. In some cases, the PAT tool 610 communicates over a network 630 with one or more computing systems 640 (*e.g.*, servers, personal computers, tablets, IoT devices, mobile phones, dedicated control units, etc.) which can execute various algorithms as described below. The computing systems 640 can also act to change one or more operating parameters associated with the bioreactor 620. In some cases, the bioreactor 620 can also have network connectivity such that it can communicate with one or more remote computing systems 640 which, in turn, can cause one or more operating parameters of the bioreactor 620 to change.

[00145] The PAT tool 610 as noted above can be used to control and monitor production processes (such as in the bioreactor 620) in-line or at-line in real time. The PAT tool 610 can utilize spectroscopic techniques such as, for example, near-infrared spectroscopy, fluorescence spectroscopy and Raman spectroscopy. Raman spectroscopy presents itself as a technology particularly suited for use in bioreactor production processes as it provides clear, sharp spectra without suffering from some of the disadvantages of other technologies such as water interference in near-infrared spectroscopy and other interferences which might make the spectra less sharp and the like. The PAT tool 610 can include a laser to implement Raman spectroscopy which is a vibrational spectroscopic technique to provide a chemical fingerprint of a substance. The PAT tool 610 in the current context is technically advantageous in that it can provide non-destructive real time measurement of a number of biotherapeutic process variables including metabolites, growth profiles, product levels, product quality attributes, nutrient feeds, culture pH, and the like.

[00146] The computing systems 640 can combine Raman spectroscopic analysis from the PAT tool 610 with chemometric modelling to provide for the real time monitoring of these variables.

The spectral peaks obtained by Raman spectroscopy are associated with one or more process variables of interest (which can be pre-defined and/or measured by offline analysis) using chemometric modelling software. Various signal processing techniques can be applied to identify and quantify the spectral peaks generated by the PAT tool 610. One type of signal processing technique is a Partial Least Squares (PLS) regression model which can be employed to model Raman data given the linear nature of the Raman signal vs. analyte concentration. In some embodiments, a linear PLS model can be employed. In certain embodiments, a nonlinear PLS model can be employed.

[00147] The PAT tool 610 can be used to measure or otherwise characterize aspects relating to large therapeutic proteins, such as mAb proteins, in the bioreactor 620. In particular, in some variations, the PAT tool 610 can be used to provide the real time monitoring of glycation and/or glycosylation profiles of therapeutic proteins in manufacturing scale bioreactors. Manufacturing scale presents many technical difficulties as opposed to smaller scale, laboratory bioreactors in that complex environmental conditions need to be addressed. The current subject matter addresses these technical difficulties as well as the intricate nature of glycoprotein formation by way of the utilized PAT tool 610 and, in some cases, the use of Raman-based PLS models.

[00148] Aspects of the PAT tool 610, the bioreactor 620, and/or the computing systems 640 herein may be realized in digital electronic circuitry, integrated circuitry, specially designed ASICs (application specific integrated circuits), computer hardware, firmware, software, and/or combinations thereof. These various implementations may include implementation in one or more computer programs that are executable and/or interpretable on a programmable system including at least one programmable processor, which may be special or general purpose, coupled to receive data and instructions from, and to transmit data and instructions to, a storage system, at least one input device, and at least one output device.

[00149] These computer programs (also known as programs, software, software applications or code) include machine instructions for a programmable processor, and, may be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine language. As used herein, the term “machine-readable medium” refers to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, solid state drive, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions

and/or data to a programmable processor, including a machine-readable medium that receives machine instructions as a machine-readable signal. The term “machine-readable signal” refers to any signal used to provide machine instructions and/or data to a programmable processor.

[00150] The computing systems 640 can include a back-end component (*e.g.*, as a data server), a middleware component (*e.g.*, an application server), or a front-end component (*e.g.*, a client computer having a graphical user interface or a Web browser through which a user may interact with an implementation of the subject matter described herein), or any combination of such back-end, middleware, or front-end components. The components of the system may be interconnected by any form or medium of digital data communication (*e.g.*, a communication network). Examples of communication networks include a local area network (“LAN”), a wide area network (“WAN”), and the Internet. The computing system 640 may include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other.

[00151] FIG. 7 is a process flow diagram 700 illustrating the determination of a glycan structure on a glycosylated molecule in which, at 710, levels are obtaining for one or more glycan structures on the glycosylated molecule, for each of a plurality of runs, using a process analytical technology (PAT) tool which obtains or otherwise generates spectral data. The levels are obtained based on processes occurring within one or more bioreactors having a volume equal to or below a first threshold. Thereafter, at 720, one or more regression models are generated based on the obtained spectral data that correlate levels of the one or more glycan structures on the glycosylated molecule with the obtained spectral data. Later, at 730, the one or more glycan structures on the glycosylated molecule are measured using the PAT tool within one or more second bioreactors having a volume equal to or above a second threshold to result in measured spectral data. At least one computing device then determines, at 740, levels of the one or more glycan structures on the glycosylated molecule within one or more bioreactors using the generated one or more regression models and based on the measured spectral data.

[00152] The one or more regression models can be refined based on the combination of the obtained spectral data and the measured spectral data.

[00153] One or more operating parameters of the one or more second bioreactors can be maintained and/or selectively modified (*i.e.*, changed) based on the determined levels to produce a desired glycosylated molecule. The one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof. The nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. The concentration of glucose in the bioreactor can be automatically modified based on the measured spectral data.

[00154] In some variations, the glycosylated molecule can be purified.

[00155] The glycan structure can take various forms including G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[00156] In some variations, the measurements can be conducted using two or more bioreactors having different volumes.

[00157] The first threshold can be of different volumes including about 25 liters or less.

[00158] The second threshold can correspond to different volumes including about 1,000 liters or more including 2,000 liters, 10,000 to about 25,000 liters, and to about 15,000 liters.

[00159] The second threshold can be at least 5X greater than the first threshold, and in some cases at least 10X greater than the first threshold, and in some cases, at least 100X greater than the first threshold, and in other cases at least 500X greater than the first threshold.

[00160] The PAT tool can utilize spectroscopic techniques including Raman spectroscopy.

[00161] The one or more regression models can include or otherwise use a partial least squares (PLS) model.

[00162] The glycosylated molecule can include a monoclonal antibody (mAb). Alternatively, the glycosylated molecule comprises a non-mAb.

[00163] The determination can be performed on-site and/or off-site. Further, the measuring and/or the determining can be performed in-line, at-line, on-line, off-line, or a combination thereof.

[00164] FIG. 8 is a process flow diagram 800 illustrating the production of a glycosylated molecule having a desired glycan structure. One or more glycan structures are measured, at 810, using a process analytical technology (PAT) tool to result in spectral data. Such measuring is within a bioreactor having a volume of equal to or above 1,000 liters. Subsequently, at 820, a level of the one or more glycan structures within the bioreactor is determined by at least one computing device using one or more regression models and based on the measured spectral data. The one or more regression models are generated using trial runs from at least one bioreactor having a volume less than or equal to 50 liters and at least one bioreactor having a volume equal to or above 1,000 liters. One or more operating parameters of the bioreactor are maintained, at 830, when the level of a desired glycan structure is above a pre-defined threshold, or the level of an undesired glycan structure is below a pre-defined threshold. Further, at 840, one or more operating parameters of the bioreactor are modified when the level of a desired glycan structures is below a pre-defined threshold, or the level of an undesired glycan structure is above a pre-defined threshold.

[00165] FIG. 9 is a process flow diagram 900 for determining glycation on a molecule. Glycation levels on the molecule are obtained, at 910, for each of a plurality of runs using a process analytical technology (PAT) tool. The obtaining can be within one or more bioreactors having a volume equal to or below a first threshold. The PAT tool can obtain or otherwise generate spectral data. One or more regression models are generated, at 920, based on the obtained spectral data that correlate glycation levels on the molecule with the obtained spectral data. Glycation on the molecule can then be measured, at 930, using the PAT tool. The measuring can be within one or more second bioreactors having a volume equal to or above a second threshold to result in measured spectral data. Later, at 940, glycation levels on the molecule within one or more bioreactors are determined by at least one computing device using the generated one or more regression models and based on the measured spectral data.

[00166] The one or more regression models can be refined based on the combination of the obtained spectral data and the measured spectral data.

[00167] One or more operating parameters of the one or more second bioreactors can be maintained and/or selectively modified (*i.e.*, changed) based on the determined levels to produce a glycation level on the molecule. The one or more operating parameters comprise a pH level, a

nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof. The nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions. The concentration of glucose in the bioreactor can be automatically modified
5 based on the measured spectral data.

[00168] In some variations, the measurements can be conducted using two or more bioreactors having different volumes.

[00169] The first threshold can be of different volumes including about 25 liters or less.

[00170] The second threshold can correspond to different volumes including about 1,000 liters
10 or more including 2,000 liters, 10,000 to about 25,000 liters, and to about 15,000 liters.

[00171] The second threshold can be at least 5X greater than the first threshold, and in some cases at least 10X greater than the first threshold, and in some cases, at least 100X greater than the first threshold, and in other cases at least 500X greater than the first threshold.

[00172] The PAT tool can utilize spectroscopic techniques including Raman spectroscopy.

[00173] The one or more regression models can include or otherwise use a partial least
15 squares (PLS) model.

[00174] The molecule can include or be a mAb. Alternatively, the molecule includes or is a non-mAb.

[00175] The determination can be performed on-site and/or off-site. Further, the measuring
20 and/or the determining can be performed in-line, at-line, on-line, off-line, or a combination thereof.

[00176] FIG. 10 is a process flow diagram 1000 illustrating production of a molecule with a desired level of glycation. Glycation on the molecule is measured, at 1010, using a process analytical technology (PAT) tool to result in spectral data. The measuring can be within a
25 bioreactor having a volume of equal to or above 1,000 liters. Later, at 1020, levels of glycation on the molecule within the bioreactor are determining, by at least one computing device using one or more regression models and based on the measured spectral data. The one or more regression models can be generated using trial runs from at least one bioreactor having a volume less than or equal to 50 liters and at least one bioreactor having a volume equal to or above 1,000

liters. One or more operating parameters of the bioreactor are maintained, at 1030, when the level of glycation on the molecule is below a pre-defined threshold. In addition, at 1040, on or more operating parameters of the bioreactor are selectively modified when the level of glycation on the molecule is above a pre-defined threshold.

5 **6. EMBODIMENTS**

[00177] A1. A method for determining a glycan structure on a glycosylated molecule, the method comprising:

obtaining, for each of a plurality of runs, levels for one or more glycan structures on the glycosylated molecule using a process analytical technology (PAT) tool, wherein the obtaining is
10 within one or more first bioreactors having a first volume equal to or below a first threshold, the PAT tool obtaining spectral data;

generating one or more regression models based on the obtained spectral data that correlate levels of the one or more glycan structures on the glycosylated molecule with the obtained spectral data;

15 measuring the one or more glycan structures on the glycosylated molecule using the PAT tool, wherein the measuring is within one or more second bioreactors having a second volume equal to or above a second threshold to result in measured spectral data; and

determining, by at least one computing device using the generated one or more regression models and based on the measured spectral data, levels of the one or more glycan structures on
20 the glycosylated molecule within one or more second bioreactors.

[00178] A2. The method of embodiment A1, further comprising refining the one or more regression models based on the combination of the obtained spectral data and the measured spectral data.

[00179] A3. The method of embodiment A1 or embodiment A2, further comprising
25 maintaining one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule.

[00180] A4. The method of embodiment A1 or embodiment A2, further comprising selectively modifying one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule.

[00181] A5. The method of embodiment A4, wherein the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof.

[00182] A6. The method of embodiment A5, wherein the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions.

[00183] A7. The method of embodiment A6, wherein the concentration of glucose is automatically modified based on the measured spectral data.

[00184] A8. The method of any one of embodiments A1 to A7, further comprising purifying the glycosylated molecule.

[00185] A9. The method of any one of embodiments A1 to A8, wherein the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[00186] A10. The method of any one of embodiments A1 to A9, wherein the obtaining is within two or more bioreactors having different volumes.

[00187] A11. The method of any one of embodiments A1 to A10, wherein the first threshold is about 250 liters or less.

[00188] A12. The method of any one of embodiments A1 to A10, wherein the first threshold is about 100 liters or less.

[00189] A13. The method of any one of embodiments A1 to A10, wherein the first threshold is about 50 liters or less.

[00190] A14. The method of any one of embodiments A1 to A10, wherein the first threshold is about 25 liters or less.

[00191] A15. The method of any one of embodiments A1 to A10, wherein the first threshold is about 10 liters or less.

[00192] A16. The method of any one of embodiments A1 to A10, wherein the first threshold is about 5 liters or less.

[00193] A17. The method of any one of embodiments A1 to A10, wherein the first threshold is about 2 liters or less.

[00194] A18. The method of any one of embodiments A1 to A10, wherein the first threshold is about 1 liter or less.

5 **[00195]** A19. The method of any one of embodiments A1 to A18, wherein the second threshold is about 1,000 liters or more.

[00196] A20. The method of any one of embodiments A1 to A18, wherein the second threshold is about 2,000 liters or more.

10 **[00197]** A21. The method of any one of embodiments A1 to A18, wherein the second threshold is about 5,000 liters or more.

[00198] A22. The method of any one of embodiments A1 to A18, wherein the second threshold is about 10,000 to about 25,000 liters.

[00199] A23. The method of any one of embodiments A1 to A18, wherein the second threshold is about 15,000 liters.

15 **[00200]** A24. The method of any one of embodiments A1 to A18, wherein the second threshold is at least 5X greater than the first threshold.

[00201] A25. The method of any one of embodiments A1 to A18, wherein the second threshold is at least 10X greater than the first threshold.

20 **[00202]** A26. The method of any one of embodiments A1 to A18, wherein the second threshold is at least 100X greater than the first threshold.

[00203] A27. The method of any one of embodiments A1 to A18, wherein the second threshold is at least 500X greater than the first threshold.

[00204] A28. The method of any one of embodiments A1 to A27, wherein the first volume is about 0.5 to about 250 liters.

25 **[00205]** A29. The method of any one of embodiments A1 to A27, wherein the first volume is about 1 to about 50 liters.

[00206] A30. The method of any one of embodiments A1 to A27, wherein the first volume is about 1 to about 25 liters.

[00207] A31. The method of any one of embodiments A1 to A27, wherein the first volume is about 1 to about 10 liters.

5 **[00208]** A32. The method of any one of embodiments A1 to A27, wherein the first volume is about 1 to about 5 liters.

[00209] A33. The method of any one of embodiments A1 to A32, wherein the second volume is about 1,000 to about 25,000 liters.

10 **[00210]** A34. The method of any one of embodiments A1 to A32, wherein the second volume is about 2,000 to about 25,000 liters.

[00211] A35. The method of any one of embodiments A1 to A32, wherein the second volume is about 5,000 to about 25,000 liters.

[00212] A36. The method of any one of embodiments A1 to A32, wherein the second volume is about 10,000 to about 25,000 liters.

15 **[00213]** A37. The method of any one of embodiments A1 to A32, wherein the second volume is about 15,000 to about 25,000 liters.

[00214] A38. The method of any one of embodiments A1 to A37, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.

20 **[00215]** A39. The method of any one of embodiments A1 to A38, wherein the one or more regression models comprise a partial least squares (PLS) model.

[00216] A40. The method of any one of embodiments A1 to A39, wherein the glycosylated molecule comprises a monoclonal antibody (mAb).

[00217] A41. The method of any one of embodiments A1 to A40, wherein the glycosylated molecule comprises a non-mAb.

25 **[00218]** A42. The method of any one of embodiments A1 to A41, wherein the determining step is performed on-site.

[00219] A43. The method of any one of embodiments A1 to A41, wherein the determining step is performed off-site.

[00220] A44. The method of any one of embodiments A1 to A43, wherein the determining step is performed in-line, at-line, on-line, off-line, or a combination thereof.

5 **[00221]** A45. The method of any one of embodiments A1 to A44, wherein the determining step is performed on-line.

[00222] A46. The method of any one of embodiments A1 to A44, wherein the determining step is performed at-line.

10 **[00223]** A47. The method of any one of embodiments A1 to A44, wherein the determining step is performed off-line.

[00224] A48. The method of any one of embodiments A1 to A44, wherein the determining step is performed in-line.

[00225] A49. The method of any of the preceding embodiments, wherein the obtaining comprises: receiving data characterizing the spectral data from the PAT tool.

15 **[00226]** A50. The method of any of the preceding embodiments, wherein the generating is executed by one or more computing devices.

[00227] B1. A method of producing a glycosylated molecule having a desired glycan structure, the method comprising:

20 measuring one or more glycan structures using a process analytical technology (PAT) tool to result in spectral data, wherein the measuring is within a bioreactor having a volume of equal to or above 1,000 liters;

determining, by at least one computing device using one or more regression models and based on the measured spectral data, a level of the one or more glycan structures within the bioreactor, wherein the one or more regression models are generated using trial runs from at least one bioreactor having a volume less than or equal to 50 liters and at least one bioreactor having a volume equal to or above 1,000 liters; and

25

maintaining one or more operating parameters of the bioreactor when:

the level of a desired glycan structure is above a pre-defined threshold, or

the level of an undesired glycan structure is below a pre-defined threshold; and
selectively modifying one or more operating parameters of the bioreactor when:

the level of a desired glycan structures is below a pre-defined threshold, or

the level of an undesired glycan structure is above a pre-defined threshold.

5 **[00228]** B2. The method of embodiment B1, wherein measuring is performed in-line, at-line, on-line, off-line, or a combination thereof.

[00229] B3. The method of embodiment B1 or embodiment B2, wherein measuring is performed in-line.

10 **[00230]** B4. The method of embodiment B1 or embodiment B2, wherein measuring is performed at-line.

[00231] B5. The method of embodiment B1 or embodiment B2, wherein measuring is performed off-line.

[00232] B6. The method of embodiment B1 or embodiment B2, wherein measuring is performed on-line.

15 **[00233]** B7. The method of any one of embodiments B1 to B6, wherein measuring occurs more than once daily.

[00234] B8. The method of any one of embodiments B1 to B6, wherein measuring occurs about every 5 to 60 minutes.

20 **[00235]** B9. The method of any one of embodiments B1 to B6, wherein measuring occurs about every 10 to 30 minutes.

[00236] B10. The method of any one of embodiments B1 to B6, wherein measuring occurs about every 10 to 20 minutes.

[00237] B11. The method of any one of embodiments B1 to B6, wherein measuring occurs about every 12.5 minutes.

25 **[00238]** B12. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 2,000 liters or more.

[00239] B13. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 5,000 liters or more.

[00240] B14. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 10,000 liters or more.

5 **[00241]** B15. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 15,000 liters or more

[00242] B16. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 10,000 liters to about 25,000 liters.

10 **[00243]** B17. The method of any one of embodiments B1 to B11, wherein the bioreactor volume is about 15,000 liters.

[00244] B18. The method of any one of embodiments B1 to B17, wherein the determining step is performed on-site.

[00245] B19. The method of any one of embodiments B1 to B17, wherein the determining step is performed off-site.

15 **[00246]** B20. The method of any one of embodiments B1 to B19, wherein the desired glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[00247] B21. The method of any one of embodiments B1 to B20, wherein bioreactor is a batch reactor, a fed-batch reactor, a perfusion reactor, or a combination thereof.

20 **[00248]** B22. The method of any one of embodiments B1 to B21, wherein the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof.

[00249] B23. The method of embodiment B22, wherein the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions.

25 **[00250]** B24. The method of embodiment B23, wherein the concentration of glucose is automatically modified based on the measured spectral data.

[00251] B25. The method of any one of embodiments B1 to B24, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.

[00252] B26. The method of any one of embodiments B1 to B25, wherein the one or more regression models comprise a partial least squares (PLS) model.

5 **[00253]** C1. A system for producing one or more glycosylated molecules comprising,

means for culturing a glycosylated molecule-producing cell line;

means for measuring a level of one or more glycan structures, wherein the means generates spectral data;

means for generating one or more regression models based on the spectral data; and

10 means for measuring a level of one or more glycan structures in a glycosylated molecule-producing cell line.

[00254] C2. The system of embodiment C1, wherein the glycosylated molecule-producing cell line is a mammalian cell line.

15 **[00255]** C3. The system of embodiment C2, wherein the mammalian cell line is a non-human cell line.

[00256] C4. The system of any one of embodiments C1 to C3, wherein the culturing comprises batch, fed-batch, perfusion, or a combination thereof.

[00257] C5. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 2,000 liters or more.

20 **[00258]** C6. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 5,000 liters or more.

[00259] C7. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 10,000 liters or more.

25 **[00260]** C8. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 15,000 liters or more.

[00261] C9. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 10,000 liters to about 25,000 liters.

[00262] C10. The system of any one of embodiments C1 to C4, wherein culturing comprises a volume of about 15,000 liters.

[00263] C11. The system of any one of embodiments C1 to C4, wherein measuring is performed in-line, at-line, on-line, off-line, or a combination thereof.

5 **[00264]** C12. The system of any one of embodiments C1 to C11, wherein measuring is performed in-line.

[00265] C13. The system of any one of embodiments C1 to C11, wherein measuring is performed on-line.

10 **[00266]** C14. The system of any one of embodiments C1 to C11, wherein measuring is performed at-line.

[00267] C15. The system of any one of embodiments C1 to C11, wherein measuring is performed off-line.

[00268] C16. The system of any one of embodiments C1 to C15, wherein measuring occurs more than once daily.

15 **[00269]** C17. The system of any one of embodiments C1 to C15, wherein measuring occurs about every 5 to 60 minutes.

[00270] C18. The system of any one of embodiments C1 to C15, wherein measuring occurs about every 10 to 30 minutes.

20 **[00271]** C19. The system of any one of embodiments C1 to C15, wherein measuring occurs about every 10 to 20 minutes.

[00272] C20. The system of any one of embodiments C1 to C15, wherein measuring occurs about every 12.5 minutes.

25 **[00273]** C21. The system of any one of embodiments C1 to C20, wherein the one or more glycan structures are selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[00274] C22. The system of any one of embodiments C1 to C21, further comprising a means for selectively modifying one or more operating parameters to enhance production of a desired one or more glycan structures.

[00275] C23. The system of embodiment C22, wherein the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof.

[00276] C24. The system of embodiment C23, wherein the nutrient level is selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions.

[00277] C25. The system of embodiment C24, wherein the concentration of glucose is automatically modified based on the spectral data.

[00278] C26. The system of any one of embodiments C1 to C25, wherein the one or more glycosylated molecules comprise a monoclonal antibody (mAb).

[00279] C27. The system of any one of embodiments C1 to C25, wherein the one or more glycosylated molecules comprise a non-mAb.

[00280] C28. The system of any one of embodiments C1 to C27, the spectral data comprises Raman spectra.

[00281] C29. The system of any one of embodiments C1 to C28, wherein the one or more regression models comprise a partial least squares (PLS) model.

[00282] C30. The system of any one of embodiments C1 to C29, further comprising a means for isolating the one or more glycosylated molecules.

[00283] D1. A system for producing one or more glycosylated molecules:

a bioreactor comprising a glycosylated molecule-producing cell line;

a process analytical technology (PAT) tool that measures one or more glycan structures and generates spectral data; and

a processor that correlates levels of one or more glycan structures with the spectral data using one or more regression models.

[00284] D2. The system of embodiment D1, wherein the bioreactor is about 2,000 liters or more.

[00285] D3. The system of embodiment D1, wherein the bioreactor is about 5,000 liters or more.

[00286] D4. The system of embodiment D1, wherein the bioreactor is about 10,000 liters or more.

5 [00287] D5. The system of embodiment D1, wherein the bioreactor is about 15,000 liters or more.

[00288] D6. The system of embodiment D1, wherein the bioreactor is about 10,000 liters to about 25,000 liters.

[00289] D7. The system of embodiment D1, wherein the bioreactor is about 15,000 liters.

10 [00290] D8. The system of any one of embodiments D1 to D7, wherein the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

[00291] D9. The system of any one of embodiments D1 to D8, wherein the glycosylated molecule-producing cell line is a mammalian cell line.

15 [00292] D10. The system of embodiment D9, wherein the mammalian cell line is a non-human cell line.

[00293] D11. The system of any one of embodiments D1 to D10, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.

20 [00294] D12. The system of any one of embodiments D1 to D11, wherein the one or more regression models comprise a partial least squares (PLS) model.

7. **EXAMPLES**

[00295] The following examples of the application are to further illustrate the nature of the application. It will be appreciated by those skilled in the art that changes could be made to the embodiments described above without departing from the broad inventive concept thereof. It is
25 understood, therefore, that this invention is not limited to the particular embodiments disclosed, but it is intended to cover modifications within the spirit and scope of the present invention as defined by the present description.

EXAMPLE 1: EXPERIMENTAL DESIGN FOR REAL TIME MONITORING OF GLYCATION AND GLYCOSYLATION

[00296] The purpose of this study was to develop PLS models of Raman spectra for real time monitoring of glycation and glycosylation (both CQAs), in a representative CHO cell culture at manufacturing scale.

[00297] Model development from reduced scale data was initially assessed. Developing model robustness was then considered by supplementing the reduced scale data with manufacturing scale data. Product quality was considered throughout the manufacturing process of the biotherapeutic mAb product, by using Raman Spectroscopy to achieve this by monitoring CQAs in real time throughout the upstream production process.

[00298] Eighteen batches of a mAb-producing CHO cell line were used in the experimental design. The stirred bioreactor scales were single use bag (SUB) 2000L (Thermo Fisher Scientific, Waltham, MA), glass 5L (Applikon Inc., Schiedam, Netherlands) and glass 1L (Eppendorf, Hamburg, Germany) and the fed-batch process lasted for 16 - 18 days. In total, nine 1L batches, seven 5L and two 2000L batches were included. Each 2000L batch was inoculated on Day 0 using a commercial seed train, ten reduced scale bioreactors were inoculated using Day 1 cell culture from a 2000L batch and six reduced scale bioreactors were inoculated on Day 0 using a laboratory seed train. Basal media and daily feed media were used for each batch executed, and process controls for all were implemented.

[00299] Two feeding strategies were employed during execution of the bioreactor runs, each consisted of two complex feeds, starting on Day 3 to each bioreactor. Twelve batches (2 x 2000L, 7 x 5L, 3 x 1L) were fed both complex feeds daily based upon a defined percentage of the vessel volume. The other 6 batches (6 x 1L) were fed the first complex feed based upon a defined percentage of the vessel and the second complex feed was delivered multiple times in a day to maintain a predefined target level of glucose in the bioreactor as part of a feed strategy study. All bioreactors were inoculated within the same seeding density limits.

[00300] Each batch had identical control targets for dissolved oxygen (DO), pH, and temperature. DO was controlled to 40% by aeration and sparged oxygen. A pH target of 6.95 was maintained using addition(s) of carbon dioxide and 2.0 M sodium carbonate. Temperature was controlled to the set point of 36.5 °C (35.5°C – 37.5°C) throughout the cell culture. The

scale-dependent process parameter agitation was transferred across scales using power per volume- calculated values. Offline samples were collected daily from each bioreactor and measured for a panel of metabolites including glucose, Lactate, titer, viable cell density and % viability using the Vi-CELL MetaFLEX (Beckman Coulter, Brea, CA), Vi-CELL XR Cell Viability Analyzer (Beckman Coulter) and Cedex Bio (Roche Holding AG, Switzerland) offline analysers. Each daily culture sample was retained and frozen at -70°C for further testing once each batch had been completed.

Glycation and Glycosylation Analysis

[00301] Glycation and glycosylation of the mAbs were characterized offline by LC/MS analysis. Briefly, protein samples for glycation analysis were first pretreated with EndoS enzyme to remove N-linked carbohydrates in order to eliminate glycoform based sample heterogeneity. The protein samples were then separated by High Performance Liquid Chromatography (HPLC) and analyzed by online electrospray ionization quadrupole time-of-flight mass spectrometry. Mass/charge data collected across the chromatographic peak was then summed and deconvoluted using Empower software (Waters Corp, Milford, MA). Detected glycation Isoforms were assigned based on deconvoluted mass spectra analysis and their relative abundance were calculated using peak intensities of centered deconvoluted mass spectra. Protein samples for glycosylation were first pretreated with 1M Dithiothreitol to separate individual mAb heavy and light chains. The protein samples were then analyzed as before by LC/MS and the glycosylation isoforms assigned and quantified as described previously.

Raman Spectral Acquisition during Bioreactor Batches

[00302] Raman spectra were gathered using two instruments for all batches with each using a multichannel Raman RXN2 system (Kaiser Optical Systems Inc., Ann Arbor, MI), which contained a 785-nm laser source and a charge-coupled device (CCD) at -40°C. The detector was connected to an MR probe, which consisted of a fiber optic excitation cable and a fiber optic collection cable (Kaiser Optical Systems Inc.). Data was collected by the MR probe attached to a bIO-Optic-220 stainless steel probe (Kaiser Optical Systems Inc.) inserted in the sterile bioreactor. The iCRaman software 4.1. (Mettler Toledo Autochem, Columbia, MD) was used to control the Raman RXN2 for spectral acquisition for all reduced scale batches (1L, 5L). The Raman Runtime HMI (Kaiser Optical Systems Inc.) was used for spectral acquisition in all

2000L batches. Collection for all Raman spectral data used the system settings of 10-s exposures for 75 scans, which resulted in a spectrum for a probe after 15 min including an overhead time of 2.5 min. Raman spectral acquisition spanned from wavenumbers 100 – 3,425 cm⁻¹. Reduced-scale vessels were protected from light interference by aluminum foil. (due to the single-use bioreactor set up in the manufacturing suite this was not required as it was enclosed in stainless steel). Intensity calibration of the instrument was performed with the Hololab Calibration Accessory (HCA) (Kaiser Optical Systems Inc.) prior to each use of the system and internal calibrations were set to occur every 24 hr throughout the bioreactor process.

EXAMPLE 2: DEVELOPMENT OF RAMAN SPECTROSCOPY BASED PLS MODELS FOR GLYCATION AND GLYCOSYLATION AT REDUCED SCALE (FLOW 1)

[00303] This example demonstrates that Raman spectroscopy based PLS models can be developed for glycation and glycosylation profiles of an exemplary therapeutic protein producing cell line during production in a bioreactor.

[00304] In this example, both Raman and offline data from the cell culture process at reduced scale (1L and 5L) was used to develop a panel of 7 chemometric PLS models for the glycation and glycosylation profiles of the mAb. Two models were considered for glycation (mono-glycated, non-glycated), and 5 models were considered for the glycosylation profile (G0F-GlcNac, G0, G0F, G1F, G2F). Chemometric modelling was performed with Simca 15.1 (Umetrics Inc., San Jose, CA).

[00305] Offline measurements for glycation and glycosylation were aligned with Raman spectra based on the time at which they were taken, beginning at Day05 in each batch. The decision to build models from this time point was based on empirical knowledge of the process and its production of mAb product at detectable levels by HPLC. All data prior to Day 05, Raman spectra and offline measurements, were excluded from model building and testing.

[00306] Each model consisted of a calibration sample set (CSS) of 15 batches (9 x 1L and 6 x 5L), for model development, and a calibration test sample set (CTSS) of 1 batch (5L), used as a blind data set for testing with the PLS model generated for each CQA. This batch was chosen as the CTSS randomly from the available 5L batch data. The X variables for each model in this flow were Raman spectra (centered) and the Y variables were Offline values for; % Mono-glycated, Non-glycated, G0F-GlcNac, G0, G0F, G1F and G2F (univariately scaled).

Wavenumber selection of the Raman spectra for all models was 415 – 1800 cm⁻¹ and 2800 – 3100 cm⁻¹. The spectral filters applied to all PLS models were Savitsky-Golay first derivative quadratic (31 cm⁻¹ point) and standard normal variate (SNV; data not shown).

[00307] Each PLS model was built and assessed for error by using a method of leave-batch-out cross validation (leave each batch out once in model development). The model error was averaged based upon prediction of the model against the omitted batch to identify the root mean square of cross-validation (RMSEcv). The RMSEcv indicates the predictive power of the model based on the data used to build the model. A lower average error (RMSEcv) indicated an improved model. This enabled better informed decision-making on which component number was to be used for the models generated for testing against the blind data set. Testing the models' predictive capability vs the CTSS, using the predict function in Simca 15.1, identified the root mean square error of prediction (RMSEP), which indicates the predictive power of the model vs an unseen dataset. The Regression (R²) value, coefficient of variation was recorded for each PLS model. This R² value is used to determine the amount of variation of the Y variable which the model predictors (X variables) can explain. The closer an R² value is to 1, the greater a model explained the Y variable. Model performance was assessed based on each models' respective RMSEcv, RMSEP and R² values. The variable importance of projections (VIPs) of each model was also considered, this is a parameter which summarizes the importance of the X-variables in a model in predicting the Y-variable. X-Variables with a value greater than 1, are considered most relevant for explaining the Y-variable.

[00308] Flow 1 evaluated the use of reduced scale data only as a calibration data set for PLS model building. Each of the 7 models developed, 2 for glycation (Mono-glycated, Non-glycated) and 5 glycosylation profiles (G0F-GlcNac, G0, G0F, G1F, G2F) were tested by predicting against a blind data set of reduced scale batch data. Each model in flow 1 consisted of 15 batches of reduced scale process data in the CSS with one full batch (5L) being used for testing as a blind CTSS.

[00309] Leave batch out cross validation was used to determine the optimum component number for each model. This was decided as the lowest number of components with the lowest matching RMSEcv value. Each model was assessed individually for optimum component number. Average R² results from cross validation show a strong degree of variability in each

model, explained by an R² value of >0.8 reported for each of the models considered (Table 1). The acceptance criteria for each model's accuracy were determined based on the range of the measurements observed in the model CSS. Accuracy acceptance criteria was set as 10% of the range of the values used in the calibration data set as per chemometric model development standards at JSI. This was determined to be an adequate acceptance criterion based on potential sources of error for each model including the acceptable error of the method employed in offline analysis. The error in each model was assessed by comparing the root mean square errors (RMSE_{cv}, RMSEP) against the acceptance criteria calculated (Table 1.). The RMSE_{cv} informs the optimum component number. This value also indicates the ability of the model to predict the variable of interest based on the calibration dataset. The component numbers chosen for each model gave an RMSE_{cv} which fell within the acceptance criteria for prediction error and as such were deemed appropriate for model development.

[00310] Table I: Flow 1: Model development statistics for Raman PLS models for glycation and glycosylation

Cross-validation data set				
CQA (%)	Component Number	R ²	RMSECV	Acceptable RMSECV
Monoglycated	6	0.8036	0.5169	≤0.554
Non-Glycated	6	0.8036	0.5169	≤0.554
G0F-GlcNac	6	0.9019	0.0312	≤0.0322
G0	5	0.9348	0.1315	≤0.1453
G0F	6	0.9546	2.1576	≤2.6964
G1F	6	0.9570	2.0707	≤2.4149
G2F	5	0.9138	0.2670	≤0.3349
Blind Prediction data set (5L)				
CQA (%)	Component Number	R ²	RMSEP	Acceptable RMSEP
Monoglycated	6	0.7131	0.5409	≤0.554
Non-Glycated	6	0.7131	0.5409	≤0.554
G0F-GlcNac	6	0.6945	0.0303	≤0.0322
G0	5	0.8777	0.1611	≤0.1453
G0F	6	0.902	1.8996	≤2.6964
G1F	6	0.9163	1.756	≤2.4149
G2F	5	0.9115	0.2273	≤0.3349

[00311] Complimentary results were observed when each model's prediction accuracy was tested with the blind data set (5L batch) (Table 1). The R^2 was >0.85 for Mono-glycated, Non-glycated, G0F, G1F, G2F, G0 and <0.7 G0F-GlcNac. All R^2 values showed a decrease when tested against the blind data set, however, the RMSEP values for all models, except for G0, mono-glycation and non-glycation, showed a lower error in prediction than the values observed for the RMSEcv. An R^2 of >0.9 does not necessarily indicate a better model and reinforces the need to consider multiple factors including RMSEcv and RMSEP in conjunction with R^2 values during model assessment.

[00312] While testing against the blind data set at reduced scale showed a decrease in R^2 and slight increase in model error observed for the G0, mono-glycation and non-glycation models, it should be noted that these models still falls within the acceptance criteria and so can be deemed suitable for use in predicting glycosylation at reduced scale. The trends observed for the prediction testing (FIG. 3A – FIG. 3G) confirm the suitability of the models for use, the Mono-glycated and Non-glycated trends follow the offline data trends tightly and are complementary to one another for the process duration as would be expected. G0F-GlcNac, G0, G0F, G1F and G2F follow the offline trends similarly and whilst slight deviations were observed in each, all were within acceptable error criteria.

[00313] Raman spectral regions identified in each model to have a VIP score > 1 indicated an acceptable degree of model specificity (FIG. 11A – FIG. 11F). Glycation models (mono-glycated, non-glycated) indicated a Raman spectral region association similar to that identified for glucose PLS models previously, while some degree of secondary correlation is likely due to the structure of glycated proteins, unwanted correlations with glucose for these models was avoided by ensuring data used in the model CSS was acquired at intervals both pre and post glucose feeding. Different glucose feeding strategies were also employed when executing bioreactor batches for data collection in order to break unwanted glucose correlations. The glycosylation profile models (G0F-GlcNac, G0, G0F, G1F, and G2F) showed VIP scores >1 in regions which have previously been associated with glycan components, for example, mannose, fucose, n-acetylglucosamine. This suggests the model identifying peaks in the Raman spectra associated with each of the glycans.

[00314] An initial assessment of the inner relation plot of each model indicated a satisfactory level of linearity to progress with model assessment, however it is possible for future work, given secondary correlations with product titer, a nonlinear-PLS method could potentially improve model accuracy. Model performance in each case is acceptable and a close agreement with daily offline samples with an expected profile supports model development decisions made in this flow.

[00315] The ability to monitor CQAs such as glycation and glycosylation in real time during a bioreactor process aligns with the key goals of QbD, with the assessment of quality being built into the process (PDA., 2012). Using in process CQA monitoring, a cause and effect relationship can be established between the critical process parameters (CPPs) and product quality. Establishing this approach early in product development when process characterization has not yet been fully completed can reduce the time needed for development and scale up (Kozlowski., 2006) as well as reducing manufacturing inefficiencies which allow for tighter control of product quality and an improvement in yield (Yu *et al.*, 2014).

[00316] The results presented here support the decisions made in model development and show that Raman spectroscopy based PLS models are capable of predicting both the glycation profiles and glycosylation, with specific focus on glycan profiles, of CHO mAb producing bioreactor processes.

EXAMPLE 3: PREDICTION OF GLYCATION AND GLYCOSYLATION AT MANUFACTURING SCALE USING MODELS DEVELOPED WITH REDUCED SCALE DATA (FLOW 2)

[00317] In this example, each of the models developed using the CSS described above in Example 2 were investigated for their capability to accurately predict at manufacturing scale. No adjustments or additional data were used to develop the models in this example.

[00318] Each model (Mono-glycated, Non-glycated, G0F-GlcNac, G0, G0F, G1F and G2F) was tested and evaluated by using the predict function in Simca 15.1 against a new CTSS comprised of data from a single manufacturing scale batch (2000L Batch A) for this cell culture process. This batch was chosen as the CTSS randomly from the available 2000L batch data. The CTSS was then used to investigate the ability of Raman models developed using only reduced scale data (1L, 5L) to predict CQAs at a manufacturing scale. A leave batch out cross validation

was not repeated as there were no changes made to the models, and as such no change to RMSECV was observed. Each of the 7 model's RMSEP and R² values were compared to the outputs of models generated in Example 2 to determine whether reduced scale data alone is sufficient for PLS model development.

- 5 **[00319]** In flow 2, model robustness was tested against manufacturing scale data (2000L). Each model from flow 1 described above in Example II was tested by predicting against a new blind CTSS comprised of process data from a 2000L batch. The model statistics for RMSECV were as observed in Flow 1 as no additional data was added or removed from the models. Average R² values of >0.80 were observed when predicted against the 2000L blind dataset in the
- 10 case of all models (Table 2) indicating good variability within each model.

[00320] Table 2: Flow 2: Model development statistics for Raman PLS models for glycation and glycosylation.

Cross-validation data set				
CQA (%)	Component Number	R ²	RMSECV	Acceptable RMSECV
Monoglycated	6	0.8036	0.5169	≤0.554
Non-Glycated	6	0.8036	0.5169	≤0.554
G0F-GlcNac	6	0.9019	0.0312	≤0.0322
G0	5	0.9348	0.1315	≤0.1453
G0F	6	0.9546	2.1576	≤2.6964
G1F	6	0.9570	2.0707	≤2.4149
G2F	5	0.9138	0.2670	≤0.3349
Blind Prediction data set				
CQA (%)	Component Number	R ²	RMSEP	Acceptable RMSEP
Monoglycated	6	0.8699	0.3553	≤0.554
Non-Glycated	6	0.8699	0.3553	≤0.554
G0F-GlcNac	6	0.8544	0.0272	≤0.0322
G0	5	0.8334	0.0997	≤0.1453
G0F	6	0.9317	2.6670	≤2.6964
G1F	6	0.9447	2.3687	≤2.4149
G2F	5	0.8916	0.4074	≤0.3349

- [00321]** The glycation models both performed well, with an RMSEP of 0.3553% observed in the case of each model (Mono-glycated and Non-glycated). Surprisingly, the RMSEP values
- 15 observed here were lower than were observed in flow 1 (described in Example 2). This would

indicate that reduced scale data alone may be sufficient to develop a model for monitoring glycation profiles of therapeutic protein (*e.g.*, mAb) processes. The interactions that lead to protein glycation are dependent on the levels of reducing sugars, temperature and time in the bioreactor (Quan *et al.*, 2008). While the specific glycation site occupancy can be difficult to control (Wei *et al.*, 2017), the overall level of glycation can be maintained at certain levels during production scale up. As such, the process variables which typically contribute to glycation are maintained at a comparable level across all bioreactor scales. For this reason, using reduced scale data may be sufficient in developing robust glycation models.

[00322] Statistics relating to the glycosylation models show a more varied response when tested against the manufacturing scale data. The RMSEP value for the G2F model was observed to be 0.4074% which was outside of the acceptance criteria and accordingly the model was deemed unsuitable for use in monitoring at manufacturing scale. Interestingly, the R^2 value for this model was >0.85 , further reinforcing the importance of consulting multiple statistics when evaluating PLS models.

[00323] The trends showing the Raman model predictions vs the offline values for the 2000L batch as it progressed (FIG. 4A – FIG. 4G) confirm the model statistics, with the G2F deviating from the offline measurement trends for the entirety of the batch (FIG. 4G). The G0F-GlcNac, G0, G0F, G1F models met the acceptance criteria and visually appear to follow the trends of the offline measurements. G0 and G0F-GlcNac performed best with a comparable or lower RMSEP value vs that obtained in flow 1. G0F and G1F showed a noticeable increase in RMSEP value and it was observed from the trends (FIG. 4E and FIG. 4F, respectively) that the predictive capability of these models begins to decline past day 07 in the process. Existing knowledge indicates that the process moves from the exponential growth phase to the stationary phase at day 07/08. As CHO cell cultures progress and move into the stationary phase of growth they reach peak antibody production. Cell size and metabolic rates change with the cells increasing in product output by as much as twice the rate in the exponential growth phase (Templeten *et al.*, 2013; Xiao *et al.*, 2017).

[00324] The models presented in flow 2 would indicate that the process changes occurring at this time point have a different effect on the glycosylation profiles at reduced scale vs manufacturing scale. Even after introducing variability into the models using spectra at multiple

scales (1L and 5L), different Raman probes/optics, different glucose feed strategies (6 x 1L) and glucose spikes (1 x 5L), it was found that the predictive capability of the models based on reduced scale data only was not robust at a manufacturing scale.

[00325] Taking the models as a combined profile, glycosylation models built using reduced scale data perform worse when predicting against manufacturing data with an increase in prediction error observed in G0, G1F and G2F. G2F, in particular, had a significant increase (79%) in prediction error putting it outside the acceptable limits for prediction error (RMSEP). While reduced scale bioreactor processes are designed to be representative of the manufacturing process, it is apparent here that the glycosylation kinetics at manufacturing scale differ to a degree whereby reduced scale data is not alone adequate for robust PLS model development.

EXAMPLE 4: INCLUSION OF MANUFACTURING SCALE DATA INTO RAMAN SPECTROSCOPY BASED PLS MODELS FOR PREDICTING GLYCATION AND GLYCOSYLATION AT MANUFACTURING SCALE (FLOW 3)

[00326] This example describes the use of an updated CSS which is comprised of the CSS data from the small scale experiments described above in Example 2, and supplemented with data from a single manufacturing scale (2000L Batch B) batch to build each PLS model. Calibration dataset for models developed in flow 1 (described above in Example 2) were supplemented with a single batch of manufacturing scale data (2000L) and prediction testing was carried out as per flow 2 (described above in Example 3), using the same 2000L blind CTSS.

[00327] The CSS of 16 batches (9 x 1L, 6 x 5L and 1 x 2000L) was then used to develop 7 updated models as per Example 2. Once completed, each of the 7 models was then assessed by leave batch out cross validation in order to derive an updated RMSEcv and component number for each model. Testing the models' predictive capability vs the new CTSS, using the predict function in Simca 15.1, identified the root mean square error of prediction (RMSEP) and the R2 value for the updated models. As in previous flows, the RMSEcv, RMSEP, R2, and VIP values output from each model were assessed to identify the models' ability to predict at manufacturing scale. They were also used to compare predictive capability against the models developed and used in flow 1 and flow 2 to determine the necessity of including manufacturing scale data in the CSS of Raman PLS models for glycation and glycosylation.

[00328] The models here (Flow 3) sought to investigate the impact of scale data on the predictive power and robustness of the models constructed in flow 1 (described above in Example 2). The models from flow 1 were supplemented with process data from a 2000L scale batch run. The addition of this 2000L batch data meant that 16 batches were used to develop each model. Cross validation was again used here to determine the optimum component number for model development and prediction testing. An average R^2 value of >0.8 was observed at cross validation for all models with each model also having an RMSEcv within the acceptable criteria (Table 3).

[00329] Table 3: Flow 3: Model development statistics for Raman PLS models for glycation and glycosylation

Cross-validation data set				
CQA (%)	Component Number	R^2	RMSECV	Acceptable RMSECV
Monoglycated	6	0.8029	0.5132	≤ 0.554
Non-Glycated	6	0.8029	0.5132	≤ 0.554
G0F-GlcNac	6	0.9007	0.0311	≤ 0.0322
G0	5	0.9312	0.1313	≤ 0.1453
G0F	6	0.9517	2.1541	≤ 2.6964
G1F	6	0.9535	2.0647	≤ 2.4149
G2F	5	0.9316	0.2669	≤ 0.3349
Blind Prediction data set (2000L)				
CQA (%)	Component Number	R^2	RMSEP	Acceptable RMSEP
Monoglycated	6	0.8992	0.3279	≤ 0.554
Non-Glycated	6	0.8992	0.3279	≤ 0.554
G0F-GlcNac	6	0.8171	0.0224	≤ 0.0322
G0	5	0.8424	0.0977	≤ 0.1453
G0F	6	0.9302	1.50925	≤ 2.6964
G1F	6	0.9438	1.43486	≤ 2.4149
G2F	5	0.9147	0.0919	≤ 0.3349

[00330] For this flow, the prediction testing was completed using the manufacturing scale blind test data set from flow 2. An average R^2 of >0.80 was observed for each model. Notably, all models showed a decrease in the RMSEP value indicating a reduced error in prediction across all models. Also the RMSEcv of each model was observed to be similar or slightly reduced for each model which would signal that the addition of manufacturing scale data to these models did

not impact the models ability to predict at small scale. The G2F model was greatly improved when the CSS was supplemented with manufacturing scale data, as evidenced by a reduction in RMSEP from 0.4074% to 0.0919% (a relative improvement in prediction error of 77.5%).

Moreover, the trends seen in FIG. 5A – FIG. 5G illustrate that the deviations observed in flow 2 (see FIG. 4A – FIG. 4G) have been accounted for and the models were able to accurately predict throughout the entire manufacturing scale process.

[00331] The regions of each model with VIP score > 1 were compared to the models created in flow 1, each model designated similar or the same regions as significant contributors to the model, in some cases the score for the regions identified changed as a result of the manufacturing scale data addition, which further supported the decisions made during model development (FIG. 11A – FIG. 11F). It should be noted that during this work models for both mono-glycation and non-glycation were developed, each model shared the same results across all three flows. As a result a single model could be created, and the complimentary result inferred. For the purpose of this work, both models were created in an effort to support the decisions made in model development and show that acceptable values were obtained in each case when testing against an unknown dataset.

[00332] Taken together, the results presented here highlight the importance of model robustness and the considerations which must be taken into account when designing model calibration data sets. Moreover, the inclusion of manufacturing scale data, while in this study only representative of 4.4% of the total data observations available for use in the CSS, results in a significant improvement in the model's predictive power and helps ensure robust model development.

EXAMPLE 5: ASSESSMENT OF THE IMPACT OF RAMAN BASED GLUCOSE FEEDBACK CONTROL ON CELL BIOREACTOR PROCESS DEVELOPMENT

[00333] The aim of this study was to investigate the impact that a continuous Raman based feedback control strategy, using a GMP ready PAT management tool, has on cell bioreactor processes (e.g., a CHO cell bioreactor process). As exemplary models, two CHO cell bioreactor processes were chosen based on their respective development stage and the development strategy. The impact of the Raman based feedback control strategy on each CHO cell bioreactor process was considered in terms of cellular growth, metabolism and productivity, as well as a

number of key process parameters and quality attributes when compared to bolus fed bioreactor processes. The results demonstrate that Raman spectroscopy is an effective PAT tool in process development and optimization.

EXAMPLE 5.1: MATERIALS AND METHODS

5 [00334] **Bioreactor Operation.** Two mAb - producing CHO cell lines (Cell Line 1, Cell Line 2) were used in the experimental design. Three bioreactor batches of Cell Line 1 (Batch A, Batch B, Batch C) and three bioreactor batches of Cell Line 2 (Batch D, Batch E, Batch F), were executed in light shielded 1L bioreactors (Eppendorf, Hamburg, Germany) with the process for each cell line lasting 14 - 15 days. Each batch was inoculated on Day 0, within the same seeding
10 density limits, from a seed train that was propagated using media and targets. Basal media and feed media were used for each batch and process controls. Cell Line 1 was used as a standard cell line for normal yield. Cell Line 2 was used as a high output cell line that has higher growth rate and resource demand.

[00335] Two feeding strategies were employed during each of the bioreactor runs. The
15 feeding strategy for Cell Line 1 consisted of a complex feed and glucose feed. Batch A was bolus fed both complex and glucose feeds once per day, starting on Day 3, based upon a defined percentage of the vessel volume. Batch B and Batch C were fed the complex feed once per day, starting on Day 3, based upon a defined percentage of the vessel and the glucose feed was automated and delivered as required, based on the Raman glucose value, to maintain a
20 predefined target level of 1g/l glucose in the bioreactor, which was determined *a posteriori*. The feeding strategy for Cell Line 2 consisted of two complex feeds and a glucose feed. Batch D was bolus fed both complex feeds and glucose feed once per day, starting on Day 3. The complex feeds were delivered using a defined percentage of the vessel volume and the glucose feed was based on achieving a defined glucose target concentration for that day. Batch E and Batch F were
25 fed both complex feeds once per day, starting on Day 3, as a defined percentage of the vessel and the glucose feed was automated and delivered as required, based on the Raman glucose value, to maintain a predefined target level of 2g/l glucose in the bioreactor, which was also determined *a posteriori*. Each batch had identical control targets for dissolved oxygen (DO), pH, and temperature. DO was controlled to 40% by aeration and sparged oxygen. A pH target of 6.95

was maintained using addition(s) of carbon dioxide and 2.0 M sodium carbonate. Temperature was controlled to the set point of 36.5 °C (35.5°C – 37.5°C) throughout the cell culture.

[00336] *Evaluation of Bioreactor Performance: Offline Sample Analysis and Online*

Parameter Trending. Offline samples were collected daily from each bioreactor and measured for glucose, lactate, titer, viable cell density and % viability using the Vi-CELL MetaFLEX (Beckman Coulter, Brea, CA), Vi-CELL XR Cell Viability Analyzer (Beckman Coulter) and Cedex Bio (Roche Holding AG, Switzerland) offline analyzers. Each daily culture sample was retained and frozen at -70°C for further testing once each batch had been completed.

[00337] Glycation was considered when evaluating the impact of glucose feedback control.

Offline samples from day 4 – day 14/15 for each of the six bioreactor batches were thawed for product quality analysis of the mAbs produced throughout the batch. Glycation of the mAbs was characterized offline by LC/MS analysis. Briefly, all samples were initially purified using protein A purification. Protein samples for glycation analysis were then pretreated with EndoS enzyme to remove N-linked carbohydrates in order to eliminate glycoform based sample heterogeneity.

The protein samples were then separated by Ultra-High Performance Liquid Chromatography (HPLC) using Waters Acquity UPLC system (Waters Corp, Milford, MA) on a reverse phase column using a gradient of acetonitrile with trifluoroacetic acid and analyzed with a Xevo G2-XS Mass Spectrometer (Waters Corp, Milford, MA) by online electrospray ionization quadrupole time-of-flight mass spectrometry. Mass/charge data collected across the chromatographic peak was then summed and deconvoluted using Masslynx (Agilent Technologies, Santa Clara, CA) or UNIFI software (Waters Corp, Milford, MA). Detected glycation isoforms were assigned based on deconvoluted mass spectra analysis and their relative abundances were calculated using peak intensities of centered deconvoluted mass spectra.

[00338] Online trends for both pH and O₂ delivery were trended and compared throughout

batch execution for both cell lines, using Dasware control software (Hamburg, Germany), as key indicators of bioreactor environment and process conditions in the bioreactor. An evaluation of both online and offline data provided indication for any process differences observed when automated feedback control of glucose was used.

[00339] *Raman Spectral Acquisition.* Raman spectra were gathered for each batch using a

multichannel Raman RXN2 system (Kaiser Optical Systems Inc., Ann Arbor, MI), which

contained a 785-nm laser source and a charge-coupled device (CCD) at -40°C . The detector was connected to an MR probe, which consisted of a fiber optic excitation cable and a fiber optic collection cable (Kaiser Optical Systems Inc.). Data was collected by the MR probe attached to a bIO-Optic-220 stainless steel probe (Kaiser Optical Systems Inc.) inserted in the sterile
5 bioreactor. The Raman Runtime HMI (Kaiser Optical Systems Inc.) was used for spectral acquisition in all batches. Collection of all Raman spectral data used the system settings of 10s exposures for 75 scans, which resulted in a spectrum for a probe after 15 min including an overhead time of 2.5 min. Raman spectral acquisition spanned from wavenumbers $100 - 3,425\text{ cm}^{-1}$. Reduced-scale vessels were protected from light interference by aluminum foil. Intensity
10 calibration of the instrument was performed with the Hololab Calibration Accessory (Kaiser Optical Systems Inc.) prior to each use of the system and internal calibrations were set to occur every 24 hr throughout the bioreactor process.

[00340] *Raman based PLS model development for Glucose.* A cell line specific Raman based glucose PLS model was developed for Cell Line 1 and Cell Line 2 in this study to facilitate
15 real time generation of glucose concentrations in the bioreactor and to facilitate the execution of real time feedback control for glucose. All chemometric modelling was performed with SIMCA 15.0 (Umetrics Inc., San Jose, CA). Offline measurements for glucose were aligned with Raman spectra based on the time at which they were taken. Each model consisted of a calibration sample set (CSS) of reduced scale and manufacturing scale data, where available. The Cell Line 1 model
20 CSS consisted of 12 batches (6 x 5L, 3 x 2000L and 3 x 1L), for model development, and a calibration test sample set (CTSS) of 1 batch (1L), was used as a blind data set for testing with the PLS model. This batch was chosen as the CTSS randomly from the available batch data at 1L scale. The Cell Line 2 model CSS consisted of 7 batches (3 x 5L and 4 x 1L), for model
25 development, and a calibration test sample set (CTSS) of 1 batch (1L), was used as a blind data set for testing with the PLS model. This batch was chosen as the CTSS randomly from the available batch data at 1L scale. The CTSS was chosen as a 1L batch as this was the scale at which both models would be deployed for real time feedback control. All model development data was collected from lab scale studies and manufacturing scale batches executed previously. For each of the PLS models generated, robustness within the CSS was ensured by including both
30 process and technology variability. A varied sampling strategy was employed in a number of batches included in the CSS in order to break any spurious correlations with batch progression.

[00341] The X variables for each model were Raman spectra (centered) and the Y variables were Offline values for glucose (univariately scaled). Wavenumber selection of the Raman spectra for each model was $415 - 1800 \text{ cm}^{-1}$ and $2800 - 3100 \text{ cm}^{-1}$. The spectral filters applied to each PLS model were Savitsky - Golay first derivative quadratic (15 cm^{-1} point) and standard normal variate (SNV; data not shown). Each PLS model was built and assessed for error by using a method of leave - batch - out cross validation (leave each batch out once in model development). The model error was averaged based upon prediction of the model against the omitted batch to identify the root mean square error of cross - validation (RMSEcv). The RMSEcv indicates the predictive power of the model based on the data used to build the model. A lower average error (RMSEcv) indicated an improved model. This enabled better informed decision - making on which component number was to be used for the models generated for testing against the blind data set. Each model was also assessed using the root mean square error of estimation (RMSEE), a model performance indicator which relates to the residuals of the data points in the CSS, i.e. the fit of the model. Testing the models' predictive capability vs the CTSS, using the predict function in Simca 15.1, identified the root mean square error of prediction (RMSEP), which indicates the predictive power of the model vs an unseen dataset. The Regression (R^2) value, coefficient of variation was recorded for each PLS model. This R^2 value is used to determine the amount of variation of the Y variable which the model predictors (X variables) can explain. The closer an R^2 value is to 1, the greater a model explained the Y variable. Model performance was assessed based on each models' respective RMSEE, RMSEcv, RMSEP and R^2 values.

[00342] **Implementation of Raman based feedback control of Glucose.** The automation of glucose feeding was facilitated by a PAT management tool, synTQ (Optimal Industrial Automation Limited, UK). Briefly, an orchestration (recipe) created within synTQ linked the Raman Instrument to the Glucose PLS model and the bioreactor system to execute feedback control. Initiation of the orchestration in synTQ signaled the Raman Instrument to begin acquiring spectral data via Open Platform Communication (OPC) connection. The spectral data generated was then sent to synTQ where it was subsequently fed into the Glucose PLS model which is contained in a SIMCA Q engine block of the orchestration. The spectral data is translated into a glucose reading for each point where spectral data is generated which was

subsequently used to calculate the glucose feed to be delivered, within the orchestration. The resulting calculation of the glucose feed was then communicated from synTQ to the Dasware bioreactor control software (Eppendorf, Germany), to start and stop the glucose feed pumps on the bioreactor system after the feed had been delivered, via OPC connection. The glucose target concentration was maintained by initiating a glucose feed to the target concentration each time the Raman and glucose model data indicated that the glucose concentration in the bioreactor had dropped below the target value. This process was repeated, for each Raman spectra generated, for the duration of the batch.

EXAMPLE 5.2: RESULTS

[00343] *Raman based PLS model development for Glucose.* Prior to execution of glucose feedback control batches, Raman based PLS models were developed using Raman spectra and offline glucose data which had been collected from previously completed batches. A cell line specific glucose model was developed for Cell Line 1 and Cell Line 2. The Cell Line 1 model contained 170 data points in its CSS and the Cell Line 2 model contained 169 data points in its CSS. A leave batch out cross validation approach was used for determining the optimum component number for each model.

[00344] FIG. 12A and FIG. 12B summarize the finalized model for each cell line, the RMSEE for the Cell Line 1 glucose model was 0.1845 g/l and was 0.3532 g/l for the Cell Line 2 glucose model. The RMSEE values coupled with R² values for both models >0.95 indicated that there was strong variability within each model and serves to support the decisions made in the data included in each model CSS. The acceptable accuracy criteria for prediction of each model was determined to be ≤ 0.5 g/l. This was determined in relation to the current glucose feed targets employed in each Cell Line bioreactor process.

[00345] For process control purposes, a glucose measurement outside of this range may have a negative influence on the process as it may result in over or underfeeding of the bioreactor. As such, the accuracy for each model was further verified in terms of the alignment of the RMSEcv and RMSEP of each glucose model to the acceptance criteria. Both the Cell Line 1 and Cell Line 2 models were determined to have an optimum component number of 6, with corresponding RMSEcv value of 0.2695 g/l and 0.3895 g/l, respectively. The component numbers chosen for each model gave an RMSEcv which fell within the acceptance criteria for prediction error

indicating the ability of the model to predict the variable of interest based on the calibration dataset and as such were deemed appropriate for model development. Complimentary RMSEP values were observed when each of the models were tested against their respective CTSS. The Cell Line 1 glucose model had an RMSEP of 0.2926 g/l and the Cell Line 2 model had an RMSEP of 0.2573 g/l. Both of these values were well within the acceptance criteria established (≤ 0.5 g/l) for model development. Good agreement between the model statistics indicated that the design considerations made in development for both models was justified and ensured each model was robust and would accurately predict and control glucose levels in each cell line's bioreactor process.

[00346] The model CSS for Cell Line 1 and 2 differ in the availability of data at the time of model creation due to differences in the stage of process development of each cell line. A manufacturing scale batch was available for inclusion in the Cell Line 1 model whereas only lab scale data was available for the Cell Line 2 model. Despite this, both models were developed with an acceptable degree of accuracy for the purpose of this work. The Cell Line 1 model may be more easily scaled up in its current state, while the Cell Line 2 model may require additional data and model adjustment. The availability of additional data collected at the manufacturing scale can improve these models for in line deployment at scale.

[00347] *Implementation of Raman based feedback control of glucose.* For this study an automated feedback control strategy which maintained glucose to a predefined setpoint concentration was considered. Two cell lines were chosen to test this strategy and each strategy was executed in duplicate and compared to the currently employed strategy for each cell line, which was a once daily bolus feed in both cell lines. Each study was executed using a single Raman instrument and as such spectral acquisition was limited to once every ~45 mins. This was determined to be an acceptable interval for automated feedback control of glucose based on previous data generated for process setpoints and glucose consumption rates for each of the cell lines considered.

[00348] The glucose trends for the Cell Line 1 batches are outlined in FIG. 13A – FIG. 13C. FIG. 13A shows the Raman glucose trend for the bolus fed batch for Cell Line 1, Batch A. The Raman measurement of glucose throughout the batch performed well with an observed RMSEP of 0.2917 g/l. The Raman measurement provides a much greater insight into the glucose trends

and consumption throughout this batch and was used to indicate whether manual delivery of glucose had sufficiently hit the target for each feed event. It was observed in Batch A that a bolus feed strategy for glucose results in a large daily spike of glucose when feeding occurs, beginning on Day 03, which would not be observed if using offline daily measurements only. This shows that there is additional process information available when using Raman as a PAT tool even in the case where it is not controlling feed delivery that would otherwise be unavailable to a process development team. The glucose concentration throughout Batch A fluctuates between 1 - 3 g/l.

[00349] In contrast, the automated feedback control batches for Cell Line 1, Batch B and Batch C, FIG. 13A and FIG. 13C show a much tighter feeding profile with the glucose concentration maintaining at 1g/l throughout each batch. The initiation of feeding began on Day 03 in each automated feedback control batch that is comparable to the timing of the initial bolus feed for Batch A. No large spikes in concentration were observed and a consistent glucose concentration was maintained for 12 days. The RMSEP for Batch B and C were 0.2887 g/l and 0.2940 g/l respectively, a low RMSEP in each case here further serves to indicate that the glucose model for Cell Line 1 was strong and performed well. A low RMSEP in each case indicated that the model was able to provide an accurate indication of the glucose concentration for the process and informed feed delivery well, within the established acceptance criteria of +/- 0.5g/L.

[00350] The Cell Line 2 glucose trends show similar results to that of Cell Line 1 and can be seen in FIG. 14A – FIG. 14C. The Batch D trend, FIG. 14A, shows the typical trajectory for the bolus fed process for Cell Line 2 with the expected large daily peaks of glucose from the single feed delivery. Additionally, Raman monitoring displays a much more information rich process which is more reflective of glucose trends when compared to single daily offline measurements. As Cell Line 2 has a greater demand for glucose the spread in concentration range is wider than Cell Line 1 with a process range of 0.5 – 4.5 g/l. The RMSEP for the glucose model in Batch D was 0.2369 g/l which was inside the acceptance criteria for performance and indicates that the data being provided by the model is accurately reflecting the actual process values. Batches E and F, FIG. 14B and FIG. 14C respectively, were controlled for 11 days to the setpoint of 2g/l.

[00351] A key difference observed in these batches was that feed control was initiated on Day 03 whereas the manual bolus fed batch began on Day 04 when the glucose level had reduced

further. Both Batch E and Batch F controlled glucose well and an RMSEP of 0.2516 g/l and 0.2119 g/l, respectively, indicates the model performance was strong in each case. It was noted that the Batch F glucose setpoint began to drift upward from the target of 2 g/l from Day 09 of the process, which was believed to be due to a calibration issue with the pump used to deliver the glucose feed. Despite this, Batch F had a tight glucose control range and only went outside acceptance criteria of +/- 0.5 g/l of the target on the final day of the process.

[00352] Raman based feedback control of glucose for both Cell Line 1 and Cell Line 2 provided the bioreactor environment with a more stable supply of glucose and prevented drastic shifts in glucose concentrations as seen in the bolus fed batches. MAb bioreactor process development and optimization is focused on the production of high product titer with a well-defined and controlled product quality profile. The mechanism of feedback control presented here is repeatable and consistent and provided a well defined feed profile for both Cell Line 1 and Cell Line 2.

[00353] PAT tools, including Raman spectroscopy present themselves as a relatively low-cost inclusion to process development outside of initial investment. Large amounts of process data are gathered very quickly in the early stages of process development by execution of high-throughput and Design of Experiments bioreactor batches. Inclusion of Raman probes for data collection in early development batches such as these allows for an intense data collection and model creation phase to occur very early which will not interfere with the tight time constraints of process development (*e.g.* mAb process development) and will allow for implementation of advanced feedback control strategies such as those presented for Cell Line 1 and Cell Line 2. As has been shown in this study, this technology can provide a wealth of additional process information as a monitoring tool, Batch A and Batch D both identified peaks and troughs in the glucose trends which would have otherwise been unnoticed when trending single daily offline measurements. Combination of this data rich approach with process control strategies can stand to change and improve the approach taken in process development and intensification.

[00354] *Evaluation of Bioreactor performance: Offline Sample Analysis and Online Parameter Trending.* The objective of this study was to deploy Raman based automated feedback control on two exemplary cell line processes and assess the impact of this feedback control on each process. Both processes were considered in terms of the impact automated

feedback control had on process variables associated with bioreactor health, productivity and environmental conditions. Any observed impacts are discussed in relation to the development stage of each process and how a PAT focused approach at this stage may support process development and intensification.

5 **[00355]** FIG. 15A – FIG. 15H show process trends for the bolus fed and automated feedback control batches for Cell Line 1. It can be observed that the growth and health of each bioreactor batch was comparable, as indicated by the VCD, viability and LDH trends in FIG. 15A, FIG. 15B and FIG. 15C. VCD trends for each batch peaked on day07/day08 between 8.00 – 8.75 $\times 10^6$ cells/ml and gradually declined for the remainder of the process. Final day viability was
10 observed between 70 – 75% and final day LDH was observed between 1400 – 1450 IU/ml for each batch. These trends and the small degree of variability observed between each batch indicated that the Raman based feedback control had no adverse impact on the Cell Line 1 process.

[00356] Variability in the bioreactor environment was observed depending on which feeding
15 strategy was employed for Cell Line 1. pH trends for Batch A, B and C, FIG. 15D, were similar until day07 when bolus fed Batch A continued to increase up to pH 7.3, whereby constant CO₂ addition was required to maintain pH within the setpoints for this batch from day07 – day12. Raman controlled Batch B and C showed trends that differ from Batch A and maintained within the setpoint range throughout the process, particularly during day07 – day12 where pH was
20 maintained at ~ pH 7.2. Online O₂ trends for Batch A, B and C differ from one another as seen in FIG. 15E. Duplicate batches B and C show a gradually increasing demand for O₂ to be delivered from day0 – day15. The O₂ demand for Batch A differs as the O₂ requirement increases and decreases throughout the bioreactor process. For both O₂ and pH, trends
25 differences begin to appear on day07, this coincides with the time at which the process reaches peak VCD. The difference in these trends observed in the bolus fed Batch A vs feedback control batches B and C would potentially indicate a difference in cellular metabolism of bolus and feedback control batches after peak VCD. The gradually increasing demand in O₂ observed throughout feedback control batches B and C along with pH trends which do not reach the upper control limits may indicate that the environment maintained in the feedback control batches is
30 more favourable in terms of cellular metabolism, potentially attributed to a more consistent delivery of glucose feeds.

[00357] The total amount of glucose feed delivered in automated feedback control batches for Cell Line 1 was less than the bolus fed batch. FIG. 15F shows a final day glucose feed volume of 51.5 ml and 54.5 ml for Batch B and C compared to a final day glucose feed volume of 59 ml for Batch A, indicating 12.7% and 7.6% less glucose feed being delivered in Batch B and C respectively. As the feedback control batches B and C are maintaining a target concentration of 1g/L automatically at every interval where a Raman measurement is obtained there is a greater level of control over glucose concentration in the bioreactor vs the bolus fed Batch A. Automated glucose delivery, here, is based on real time measurements which capture the demand for glucose in the bioreactor more accurately than a single daily offline sample. The bolus delivery of glucose is based on feed targets which have been established from previous batch execution and by measuring the glucose concentration offline. This approach does not take into account the current performance of the bioreactor and can potentially lead to an overestimation of the glucose requirement, as identified here in Batch A.

[00358] Automated feedback control for Cell Line 1 resulted in a comparable product titer when compared to the bolus feed strategy, FIG. 15G shows a final day titer between 2.2 – 2.5 g/L for each batch as expected. A reduction of overall product glycation in both Batch B and Batch C was observed when compared to the bolus fed Batch A. Final day glycation for Batch A was observed at 6.68% while Batch B was observed at 3.78% and Batch C at 2.81%, representing a 43.4% and 57.9% reduction in glycation respectively, as seen in FIG. 15H. A comparable final day titer in all 3 Cell Line 1 batches further serves to support that no negative impact is observed on this process when Raman based feedback control is used. What is notable is the marked improvement in product glycation, this can be attributed to the lower total glucose volume in both Batch B and C as a result of Raman based feedback control. The overall concentration of glucose in a bioreactor directly impacts the level of glycation and control of this CQA identifies a key process improvement associated with the PAT approach that could otherwise not be achieved during process development.

[00359] A similar assessment of process variation was considered for the bolus fed and automated feedback control batches for Cell Line 2. FIG. 16A – FIG. 16H summarize the comparison of each batch. VCD, Viability and LDH trends indicate differences in the growth profiles of Batch D, E and F. FIG. 16A shows both the bolus fed batch D and automated feedback control batch E reached peak VCD of $10 - 10.5 \times 10^6$ cells/ml on day08 of the process

while the automated feedback control batch F peaked on day10 at 11.97×10^6 cells/ml. The drop off in VCD was lower in the automated feedback control batches E and F than in the bolus fed batch D, with a final day VCD of 4.81×10^6 cells/ml and 8.07×10^6 cells/ml in batch E and F, respectively, compared to 3.65×10^6 cells/ml in batch D. The viability and LDH trends further highlight differences in bioreactor health as seen in FIG. 16B and FIG. 16C. Bioreactor viability remained higher in both automated feedback control batches than the bolus fed batch, with batch E (56%) and F (72.5%) showing 11.1% and 27.6% greater viability than batch D (44.9%) on the final day of the process. LDH was observed to be higher in the bolus fed batch from day08 through to the day14 than the automated feedback control batches with batch E (6053 IU/ml) and F (4834.5 IU/ml) observed to be 21% and 37% lower than batch D (7694 IU/ml) on day14. A less severe drop off in both VCD and viability for Batch E and F along with a lower LDH would indicate that a Raman based feedback control strategy to a glucose setpoint of 2g/L has a positive effect on cellular growth and health in the bioreactor. A steadier glucose concentration in the bioreactor environment appears to be much more favorable for the Cell Line 2 process than the daily peaks and troughs associated with bolus feeding as seen in Batch D. An ability to keep a process alive and generating product output for longer serves as another justification for inclusion of PAT at the process development stage as the same prolonged level of growth may not be attainable with traditional development strategies.

[00360] Bioreactor pH, O₂ flow and glucose feed volume were tracked online up to day13 for each process, a connection issue prevented the remainder of the process data from being recorded in each case. It was observed that the pH trend for each batch was similar, FIG. 16D, showing daily spikes up to pH 7.1 which coincided with daily addition of the complex feeds. FIG. 16E shows that O₂ demand was also consistent across all 3 batches, each with a gradually increasing demand for O₂ up until day10. From day10 onward, the batch requirements for O₂ differed as both batch E and F O₂ flow maintains at >0.5 sl/h while Batch D O₂ flow gradually declines to below 0.5 sl/h. FIG. 15F shows a similar final day glucose feed volume for Batch D, E and F of ~26mls. Consistency in these parameters indicates no negative impact associated with Raman based feedback control for the Cell Line 2 process. The drop off in O₂ demand in the bolus Batch D is likely due to the lower amount of viable cell in this batch from day10 onward. Interestingly the total glucose delivered in both the bolus fed Batch D and Raman based feedback control batches E and F were comparable which would indicate the previous observed

impact to bioreactor health was not due to a lesser amount of feed delivery but instead the delivery system of the glucose feed, single daily bolus vs multiple smaller feeds a day.

[00361] Productivity of the Cell Line 2 process was increased when automated feedback control of glucose was used. FIG. 16G shows that the final day titer for batch E (5.75 g/L) and batch F (5.77 g/L) were 25% higher than the bolus fed batch D (4.6 g/L). The quality of the product titer was also improved in the automated feedback control batches, FIG. 16H, where final day product glycation of the bolus fed batch D was observed to be 3.06% while batch E and F were observed to be lower at 2.13% and 2.37% respectively. While all three batches show a similar total volume of glucose added, the glycation results would indicate that continuous control to a setpoint concentration is a more favorable strategy in terms of product quality. The improvements in growth, product output and quality show that adoption of PAT development strategies for a more intensified process such as that of Cell Line 2 can drive development further and produce results that may be out of reach with a traditional bioreactor process.

[00362] Once daily, bolus delivery of glucose as a platform feeding strategy is widely considered in mAb bioreactor process development as it has been shown to adequately deliver nutrients to the bioreactor to facilitate growth and product output. A primary concern of this approach to nutrient delivery is the impact on product quality. Large daily swings in glucose concentration can be observed in bolus fed batches in this study, FIG. 13A and FIG. 14A. The influence of glucose on product quality is seen primarily in its promotion of product glycation. Glycation is a critical quality attribute of mAb bioreactor processes, the direct impact of glycation on mAbs is varied and its impact is most likely dependent on the mAb sites which are glycosylated and the overall amount of glycation present.

[00363] Automation of this process removes the extra resource requirements of a more frequent bioreactor feeding schedule and adds a QbD consideration to process development which can allow for further improvements to be made, using techniques that complement current development methods. The automated feedback control to a setpoint glucose concentration using a Raman based PLS model as shown in this study had a direct positive influence on each cell line process without any additional resource requirements. Automated feedback control for Cell Line 1 resulted in a reduction of overall product glycation in both Batch B and Batch C when compared to the bolus fed Batch A, while maintaining a comparable growth profile and product

output. This improvement can be attributed to the consistent delivery and overall reduced volume of glucose feed required for the automated feedback control batches B and C. Cell Line 1 is representative of a platform process in late stage development. While significant process development has been completed for this cell line, introduction of PAT tools such as Raman based automated feedback control could stand to benefit process development without any expected negative impact on the process designed up to this point.

[00364] Cell Line 2 is representative of a more intensified process in an earlier stage of development than Cell Line 1. Process intensification in mAb producing bioreactors is mainly dependent on the biological limit of the process. Media formulation and enrichment as well as feeding strategy optimization are important to maintain higher VCD, cellular metabolism and productivity of intensified processes. The automation of glucose feeding for Cell Line 2 allowed for a more optimized feeding strategy to be tested. While total volume of glucose delivered was similar in both bolus and automated feedback control batches, a number of process improvements were observed. An improved growth and viability profile coupled with a greater product output and an improved product quality in both automated feedback control batches for Cell Line 2 highlights how effective PAT tools can be in the development process. The ability to directly influence product quality and output in process development is aligned with the goals of the QbD initiative.

[00365] Implementation of PAT technology such as Raman spectroscopy at early or late stage development can have a significant impact on process development, scale up and the robustness of manufacturing processes. Here, two CHO cell line bioreactor processes saw an indication of improvement in the efficiency of their current process when employing PAT methods at the development stage.

[00366] Employment of methods such as automated feedback control with Raman in the development phase is able to generate valuable data, and there is low risk involved. The ability to standardize PAT in process development is able to lead to a more efficient and cost-effective manufacturing process. Increased product output and/or improved product quality at manufacturing scale, as a direct result of more efficient development, may influence the dosage requirements for a patient (which could lead to more doses of a product generated per batch resulting in major cost and time savings. The consideration of technology such as this at

development and implementation in commercial manufacturing creates an end to end QbD approach which will maximize process efficiency and meet the demands for novel mAb therapies.

[00367] The results provided here demonstrated inclusion of Raman spectroscopy as a PAT tool in process development can improve product output and reduce overall product glycation. For example, the Cell Line 1 process was improved by reducing total glucose delivery to the bioreactor and in turn reducing the overall product glycation by 43.7% and 57.9% in the Raman based feedback control batches while maintaining comparable growth profiles and product output. In addition, the Cell Line 2 process was also improved with a prolonged cellular health and greatly reduced drop off in cellular viability as the process progressed. This improvement had a knock-on effect of increasing product output for this process by up to 25%. Thus, inclusion of Raman spectroscopy as a PAT tool in process development can have a significant impact on process development, scale up and the robustness of manufacturing processes.

* * *

[00368] In the descriptions above and in the claims, phrases such as “at least one of” or “one or more of” may occur followed by a conjunctive list of elements or features. The term “and/or” may also occur in a list of two or more elements or features. Unless otherwise implicitly or explicitly contradicted by the context in which it used, such a phrase is intended to mean any of the listed elements or features individually or any of the recited elements or features in combination with any of the other recited elements or features. For example, the phrases “at least one of A and B;” “one or more of A and B;” and “A and/or B” are each intended to mean “A alone, B alone, or A and B together.” A similar interpretation is also intended for lists including three or more items. For example, the phrases “at least one of A, B, and C;” “one or more of A, B, and C;” and “A, B, and/or C” are each intended to mean “A alone, B alone, C alone, A and B together, A and C together, B and C together, or A and B and C together.” Use of the term “based on,” above and in the claims is intended to mean, “based at least in part on,” such that an unrecited feature or element is also permissible.

[00369] The subject matter described herein can be embodied in systems, apparatus, methods, and/or articles depending on the desired configuration. The implementations set forth in the foregoing description do not represent all implementations consistent with the subject matter

described herein. Instead, they are merely some examples consistent with aspects related to the described subject matter. Although a few variations have been described in detail above, other modifications or additions are possible. In particular, further features and/or variations can be provided in addition to those set forth herein. For example, the implementations described above
5 can be directed to various combinations and subcombinations of the disclosed features and/or combinations and subcombinations of several further features disclosed above. In addition, the logic flows depicted in the accompanying figures and/or described herein do not necessarily require the particular order shown, or sequential order, to achieve desirable results. Other implementations can be within the scope of the following claims.

What is claimed is:

1. A method for determining a glycan structure on a glycosylated molecule, the method comprising:

obtaining, for each of a plurality of runs, levels for one or more glycan structures on the glycosylated molecule using a process analytical technology (PAT) tool, wherein the obtaining is within one or more first bioreactors having a first volume equal to or below a first threshold, the PAT tool obtaining spectral data;

generating one or more regression models based on the obtained spectral data that correlate levels of the one or more glycan structures on the glycosylated molecule with the obtained spectral data;

measuring the one or more glycan structures on the glycosylated molecule using the PAT tool, wherein the measuring is within one or more second bioreactors having a second volume equal to or above a second threshold to result in measured spectral data; and

determining, by at least one computing device using the generated one or more regression models and based on the measured spectral data, levels of the one or more glycan structures on the glycosylated molecule within one or more second bioreactors.

2. The method of claim 1, further comprising refining the one or more regression models based on the combination of the obtained spectral data and the measured spectral data.

3. The method of claim 1 or 2, further comprising maintaining one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule.

4. The method of claim 1 or 2, further comprising selectively modifying one or more operating parameters of the one or more second bioreactors based on the determined levels to produce a desired glycosylated molecule,

wherein the one or more operating parameters optionally comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof,

wherein the nutrient level is optionally selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions, and

wherein the concentration of glucose is optionally automatically modified based on the measured spectral data.

5. The method of any one of claims 1 to 4, further comprising purifying the glycosylated molecule.
6. The method of any one of claims 1 to 5, wherein the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.
7. The method of any one of claims 1 to 6, wherein the obtaining is within two or more bioreactors having different volumes.
8. The method of any one of claims 1 to 7, wherein
 - (a) the first threshold is:
 - (i) about 250 liters or less;
 - (ii) about 100 liters or less;
 - (iii) about 50 liters or less;
 - (iv) about 25 liters or less;
 - (v) about 10 liters or less;
 - (vi) about 5 liters or less;
 - (vii) about 2 liters or less;
 - (viii) about 1 liter or less; and/or

- (b) the second threshold is:
 - (i) about 1,000 liters or more;
 - (ii) about 2,000 liters or more;
 - (iii) about 5,000 liters or more;
 - (iv) about 10,000 liters or more;
 - (v) about 15,000 liters or more;
 - (vi) about 10,000 to about 25,000 liters
 - (vii) at least 5X greater than the first threshold;
 - (viii) at least 10X greater than the first threshold;
 - (ix) at least 100X greater than the first threshold; or
 - (x) at least 500X greater than the first threshold.
- 9. The method of any one of claims 1 to 8, wherein
 - (a) the first volume is:
 - (i) about 0.5 to about 250 liters;
 - (ii) about 1 to about 50 liters;
 - (iii) about 1 to about 25 liters;
 - (iv) about 1 to about 10 liters; or
 - (v) about 1 to about 5 liters; and/or
 - (b) the second volume is:
 - (i) about 1,000 to about 25,000 liters;

- (ii) about 2,000 to about 25,000 liters;
- (iii) about 5,000 to about 25,000 liters;
- (iv) about 10,000 to about 25,000 liters; or
- (v) about 15,000 to about 25,000 liters.

10. The method of any one of claims 1 to 9, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.

11. The method of any one of claims 1 to 10, wherein the one or more regression models comprise a partial least squares (PLS) model.

12. The method of any one of claims 1 to 11, wherein the glycosylated molecule comprises a monoclonal antibody (mAb) or a non-mAb.

13. The method of any one of claims 1 to 12, wherein the determining step is performed

- (a) on-site or off-site; and/or
- (b) in-line, at-line, on-line, off-line, or a combination thereof.

14. The method of any of the preceding claims, wherein the obtaining comprises: receiving data characterizing the spectral data from the PAT tool.

15. The method of any of the preceding claims, wherein the generating is executed by one or more computing devices.

16. A method of producing a glycosylated molecule having a desired glycan structure, the method comprising:

measuring one or more glycan structures using a process analytical technology (PAT) tool to result in spectral data, wherein the measuring is within a bioreactor having a volume of equal to or above 1,000 liters;

determining, by at least one computing device using one or more regression models and based on the measured spectral data, a level of the one or more glycan structures within the

bioreactor, wherein the one or more regression models are generated using trial runs from at least one bioreactor having a volume less than or equal to 50 liters and at least one bioreactor having a volume equal to or above 1,000 liters; and

maintaining one or more operating parameters of the bioreactor when:

the level of a desired glycan structure is above a pre-defined threshold, or

the level of an undesired glycan structure is below a pre-defined threshold; and

selectively modifying one or more operating parameters of the bioreactor when:

the level of a desired glycan structures is below a pre-defined threshold, or

the level of an undesired glycan structure is above a pre-defined threshold.

17. The method of claim 16, wherein measuring is performed in-line, at-line, on-line, off-line, or a combination thereof.

18. The method of claim 16 or 17, wherein measuring occurs:

- (a) more than once daily;
- (b) about every 5 to 60 minutes;
- (c) about every 10 to 30 minutes;
- (d) about every 10 to 20 minutes; or
- (e) about every 12.5 minutes.

19. The method of any one of claims 16 to 18, wherein measuring is within a bioreactor having a volume of:

- (a) about 2,000 liters or more;
- (b) about 5,000 liters or more;
- (c) about 10,000 liters or more;
- (d) about 15,000 liters or more;
- (e) about 10,000 liters to about 25,000 liters; or

- (f) about 15,000 liters.
20. The method of any one of claims 16 to 19, wherein the determining step is performed on-site or off-site.
21. The method of any one of claims 16 to 20, wherein the desired glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.
22. The method of any one of claims 16 to 21, wherein the bioreactor is a batch reactor, a fed-batch reactor, a perfusion reactor, or a combination thereof.
23. The method of any one of claims 16 to 22, wherein the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof,
- wherein the nutrient level is optionally selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions; and
- wherein the concentration of glucose is optionally automatically modified based on the measured spectral data.
24. The method of any one of claims 16 to 23, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.
25. The method of any one of claims 16 to 24, wherein the one or more regression models comprise a partial least squares (PLS) model.
26. A system for producing one or more glycosylated molecules comprising,
means for culturing a glycosylated molecule-producing cell line;
means for measuring a level of one or more glycan structures, wherein the means generates spectral data;
means for generating one or more regression models based on the spectral data; and

means for measuring a level of one or more glycan structures in a glycosylated molecule-producing cell line.

27. The system of claim 26, wherein the glycosylated molecule-producing cell line is a mammalian cell line, optionally wherein the mammalian cell line is a non-human cell line.

28. The system of claim 26 or 27, wherein the culturing comprises batch, fed-batch, perfusion, of a combination thereof.

29. The system of any one of claims 26 to 28, wherein culturing comprises a volume of:

- (a) about 2,000 liters or more;
- (b) about 5,000 liters or more;
- (c) about 10,000 liters or more;
- (d) about 15,000 liters or more;
- (e) about 10,000 liters to about 25,000 liters; or
- (f) about 15,000 liters.

30. The system of any one of claims 26 to 29, wherein measuring is performed in-line, at-line, on-line, off-line, or a combination thereof.

31. The system of any one of claims 26 to 30, wherein measuring occurs:

- (a) more than once daily;
- (b) about every 5 to 60 minutes;
- (c) about every 10 to 30 minutes;
- (d) about every 10 to 20 minutes; or
- (e) about every 12.5 minutes.

32. The system of any one of claims 26 to 31, wherein the one or more glycan structures are selected from the group consisting of: G0F-GlcNAc, G0, G0F, G1F, and G2F, or a combination thereof.

33. The system of any one of claims 26 to 32, further comprising a means for selectively modifying one or more operating parameters to enhance production of a desired one or more glycan structures,

wherein the one or more operating parameters comprise a pH level, a nutrient level, a culture media concentration, a frequency interval for culture media addition, or a combination thereof,

wherein the nutrient level is optionally selected from the group consisting of: a concentration of glucose, a concentration of lactate, a concentration of glutamine, and a concentration of ammonium ions, and

wherein the concentration of glucose is optionally automatically modified based on the measured spectral data.

34. The system of any one of claims 26 to 33, wherein the one or more glycosylated molecules comprise a monoclonal antibody (mAb) or a non-mAb.

35. The system of any one of claims 26 to 34, wherein the spectral data comprises Raman spectra.

36. The system of any one of claims 26 to 35, wherein the one or more regression models comprise a partial least squares (PLS) model.

37. The system of any one of claims 26 to 36, further comprising a means for isolating the one or more glycosylated molecules.

38. A system for producing one or more glycosylated molecules:
a bioreactor comprising a glycosylated molecule-producing cell line;

a process analytical technology (PAT) tool that measures one or more glycan structures and generates spectral data; and

a processor that correlates levels of one or more glycan structures with the spectral data using one or more regression models.

39. The system of claim 38, wherein the bioreactor is:

- (a) about 2,000 liters or more;
- (b) about 5,000 liters or more;
- (c) about 10,000 liters or more;
- (d) about 15,000 liters or more;
- (e) about 10,000 liters to about 25,000 liters; or
- (f) about 15,000 liters.

40. The system of claim 38 or 39, wherein the glycan structure is selected from the group consisting of: G0F-GlcNac, G0, G0F, G1F, and G2F, or a combination thereof.

41. The system of any one of claims 38 to 40, wherein the glycosylated molecule-producing cell line is a mammalian cell line, optionally wherein the mammalian cell line is a non-human cell line.

42. The system of any one of claims 38 to 41, wherein the PAT tool utilizes or otherwise comprises Raman spectroscopy.

43. The system of any one of claims 38 to 42, wherein the one or more regression models comprise a partial least squares (PLS) model.

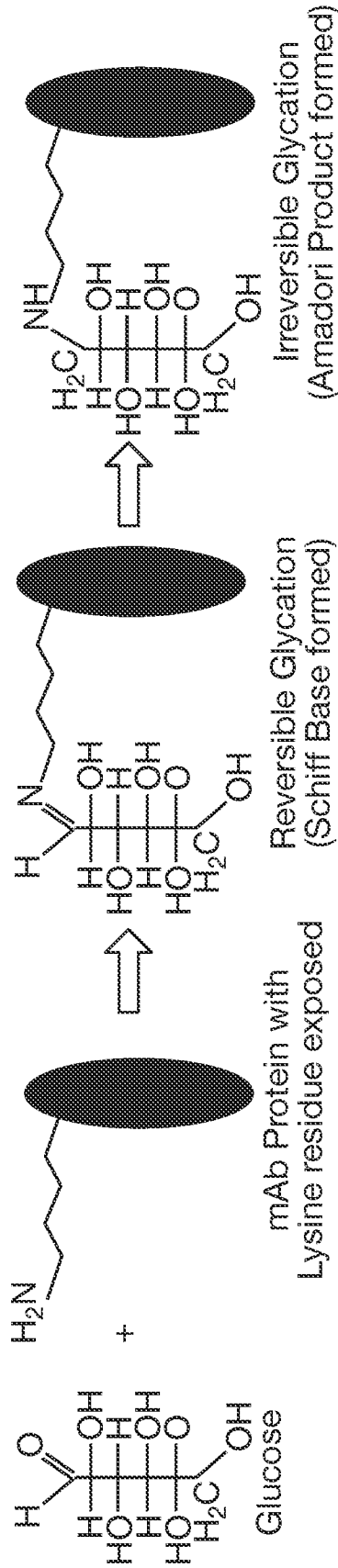


FIG. 1A

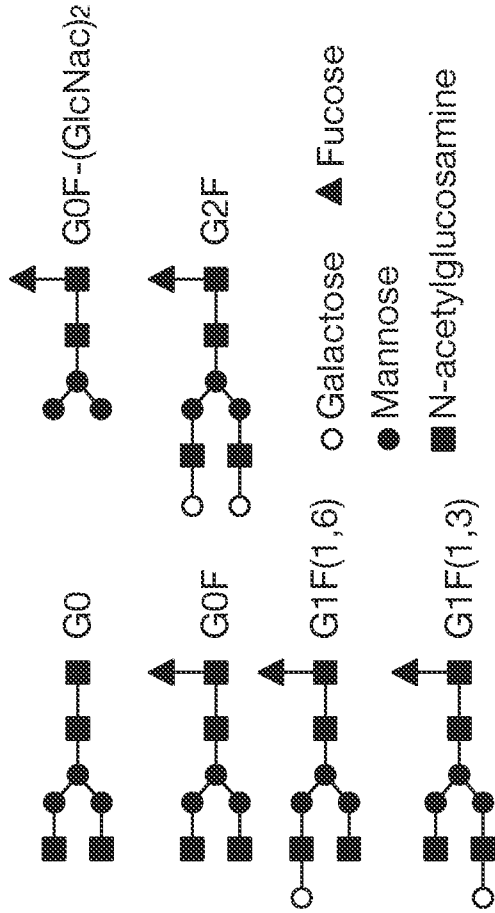
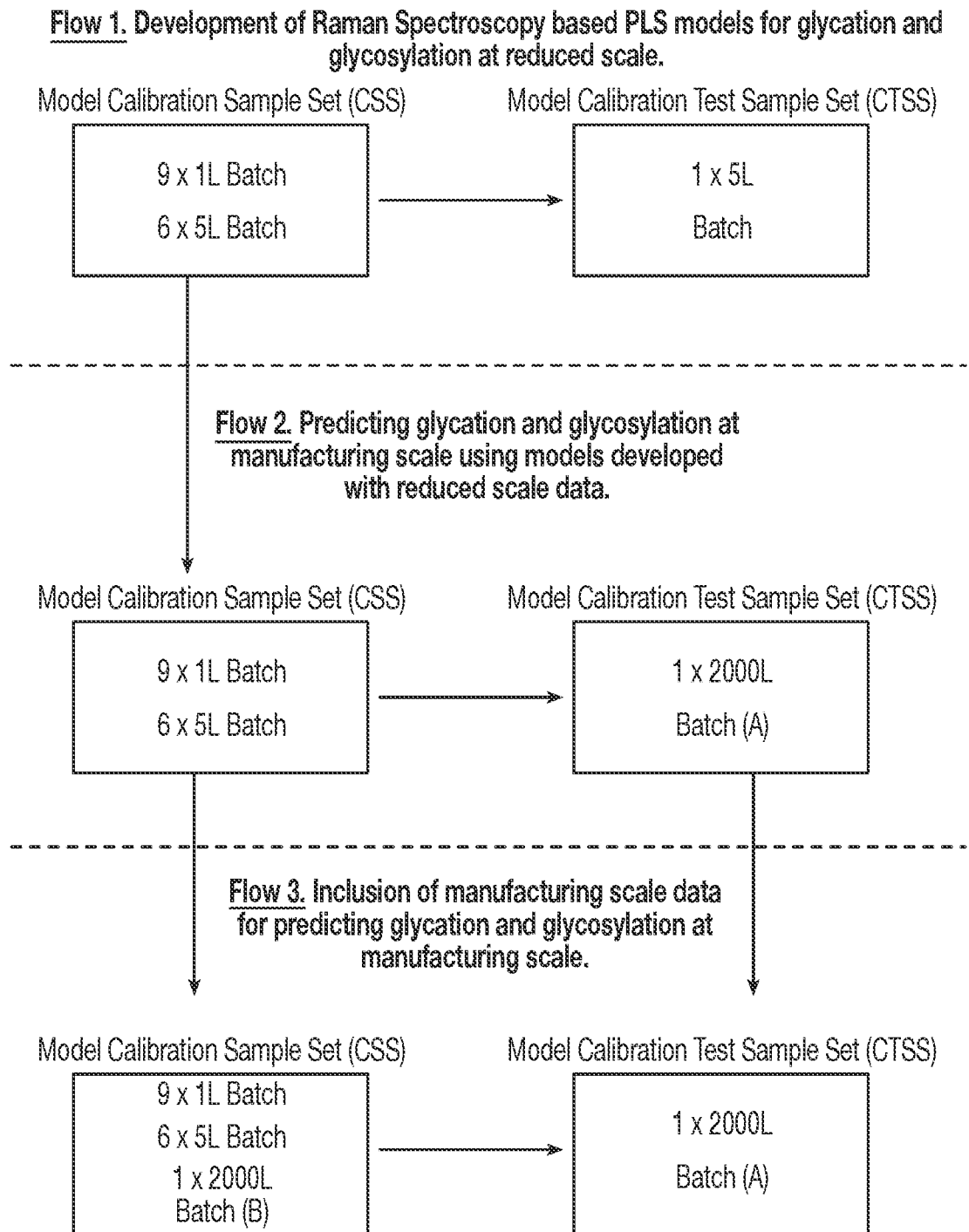


FIG. 1B

2/14

**FIG. 2**

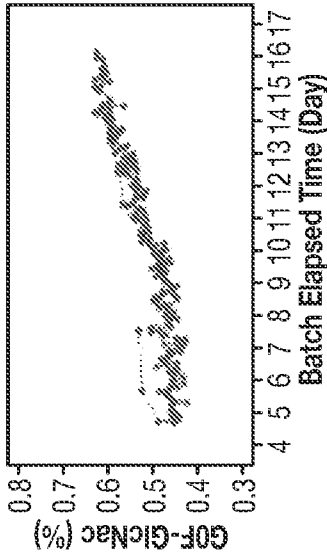


FIG. 3A

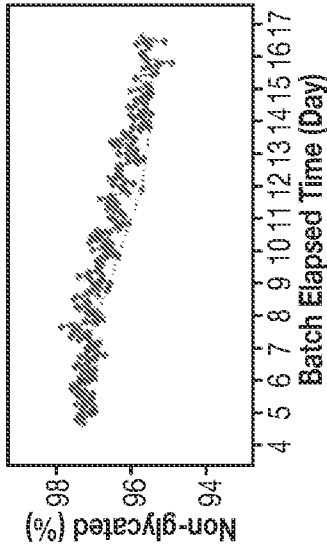


FIG. 3B

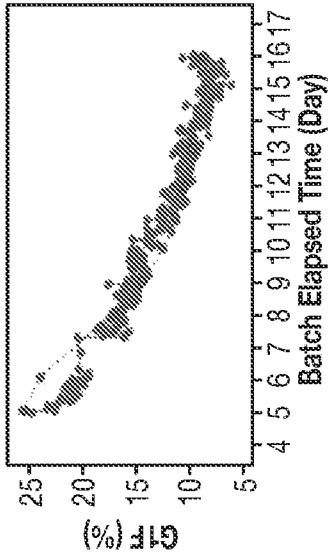


FIG. 3C

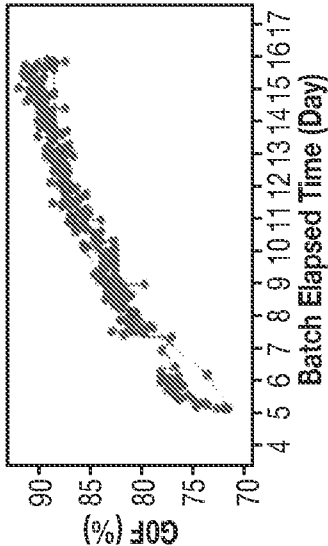


FIG. 3D



FIG. 3E

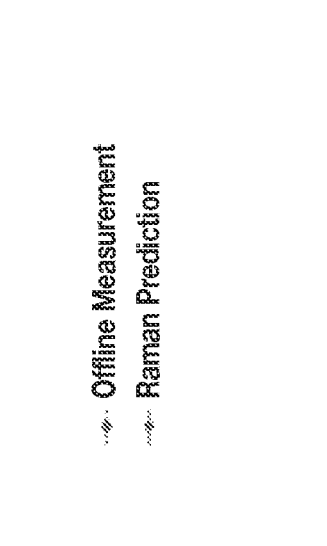


FIG. 3F

Offline Measurement
Raman Prediction

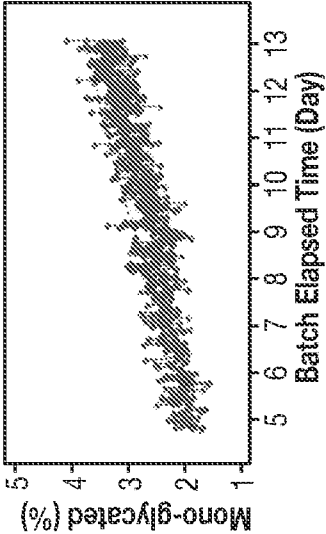


FIG. 4A

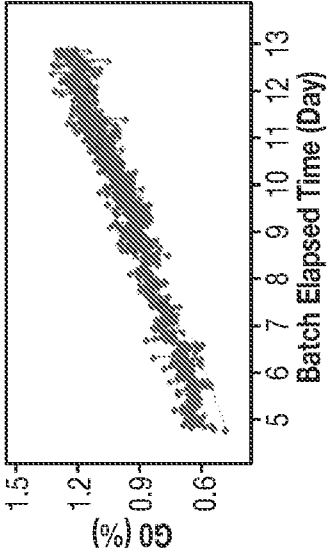


FIG. 4D

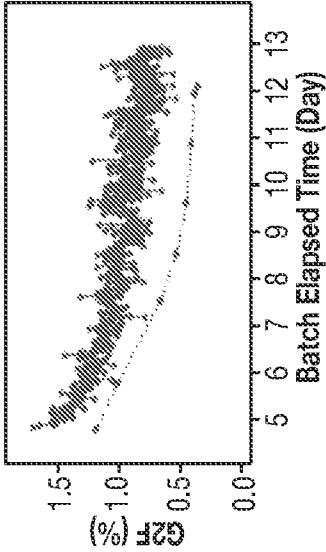


FIG. 4G

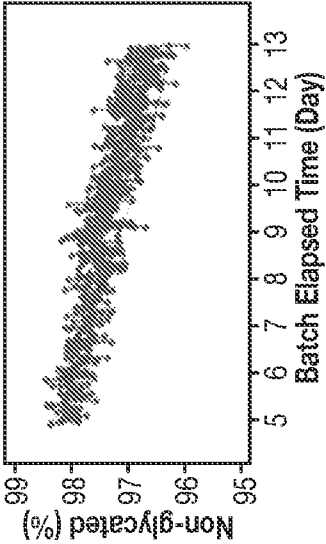


FIG. 4B

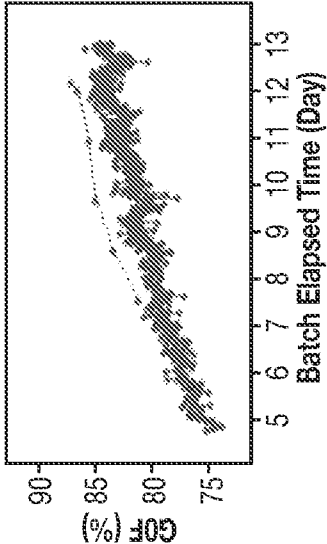


FIG. 4E

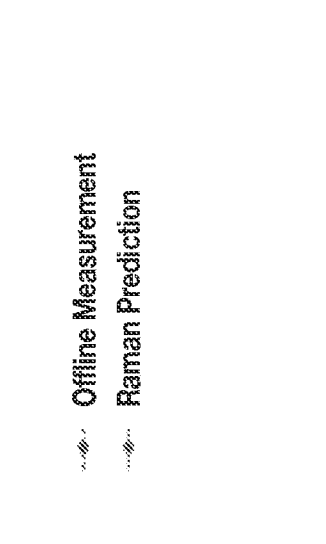


FIG. 4F

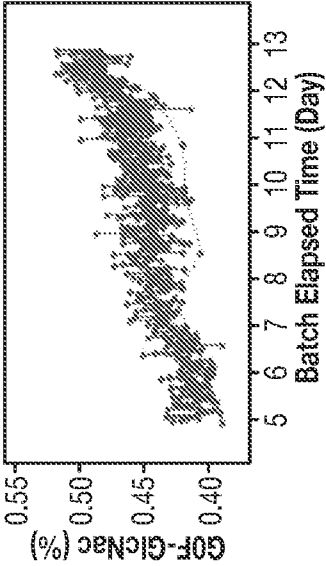


FIG. 4C

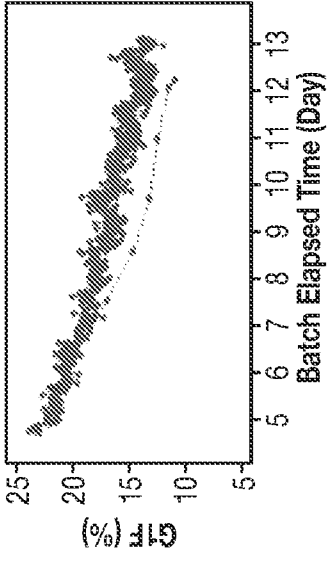


FIG. 4H

Offline Measurement
Raman Prediction

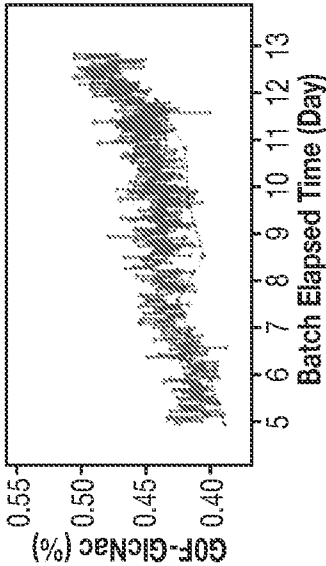


FIG. 5C

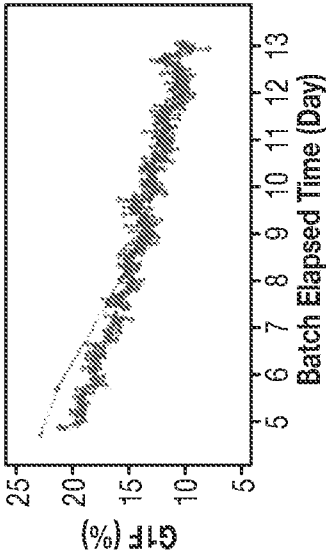


FIG. 5F

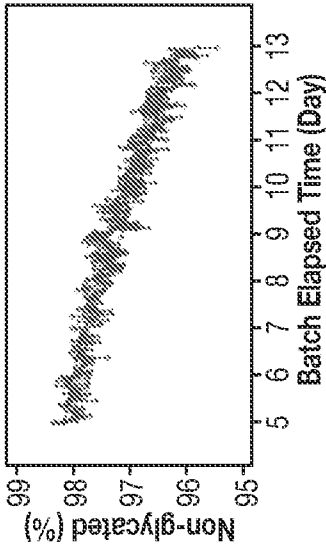


FIG. 5B

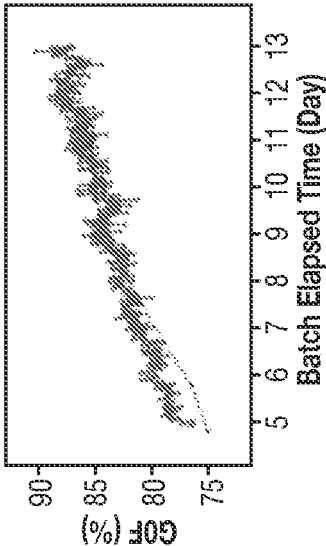


FIG. 5E

Offline Measurement
Raman Prediction

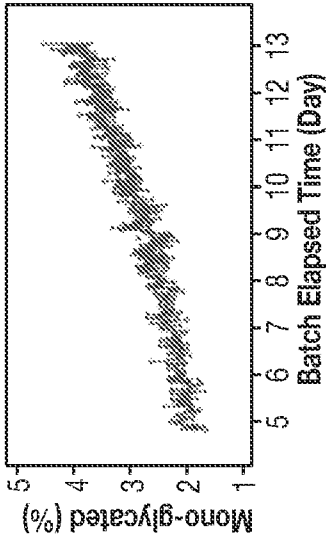


FIG. 5A

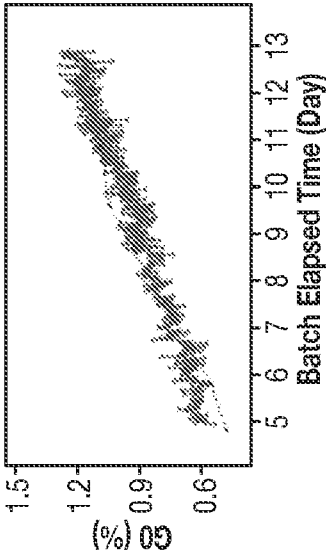


FIG. 5D

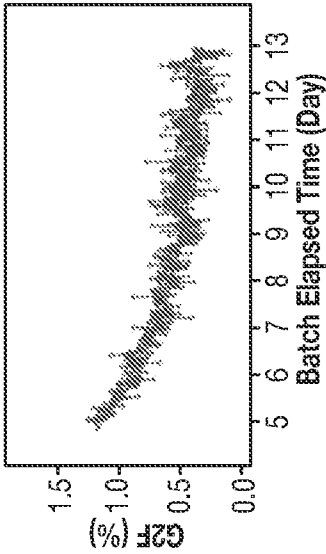
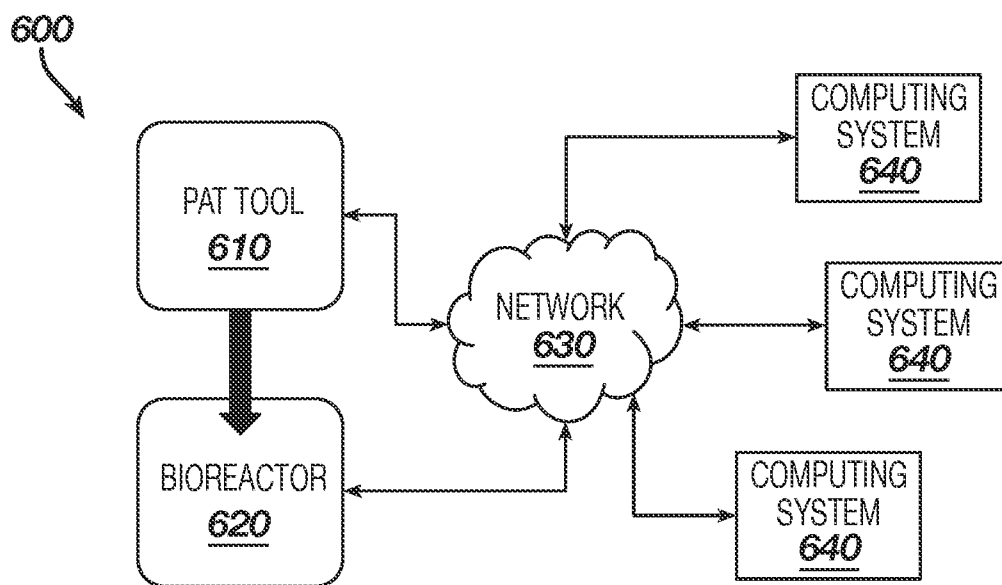
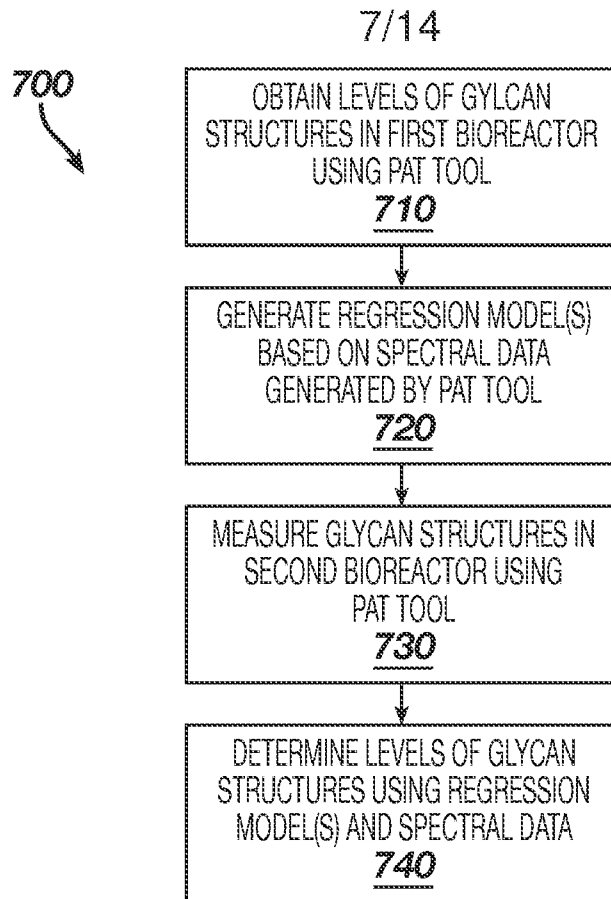
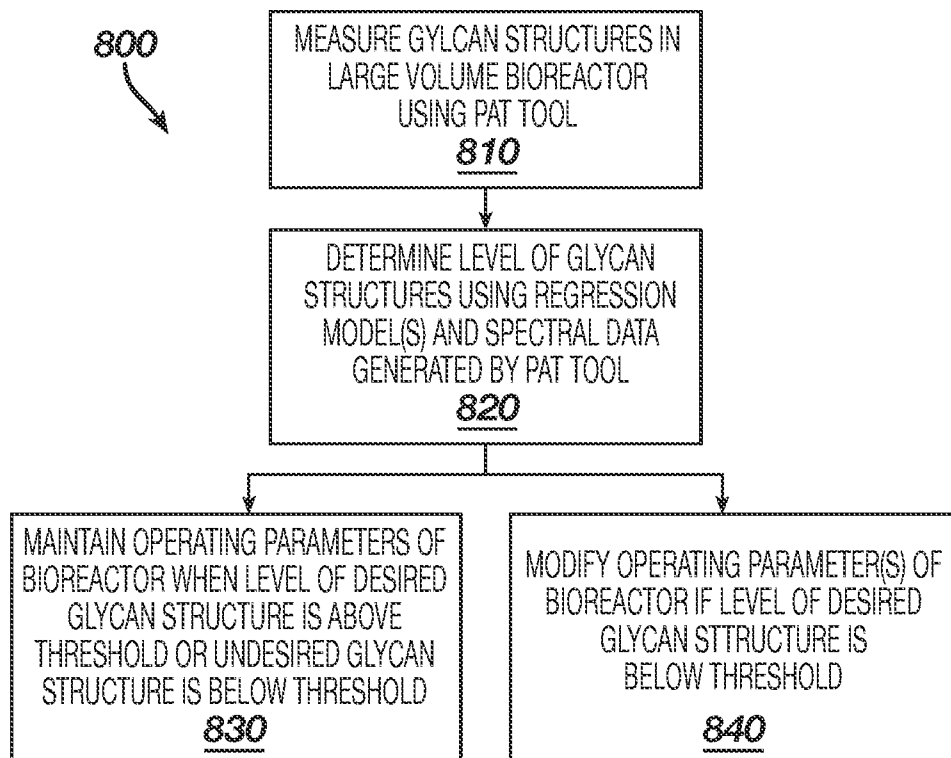


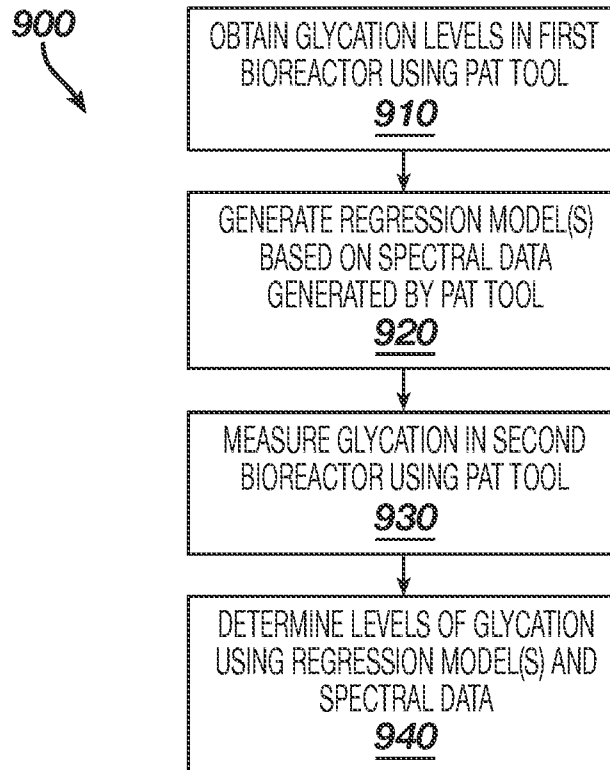
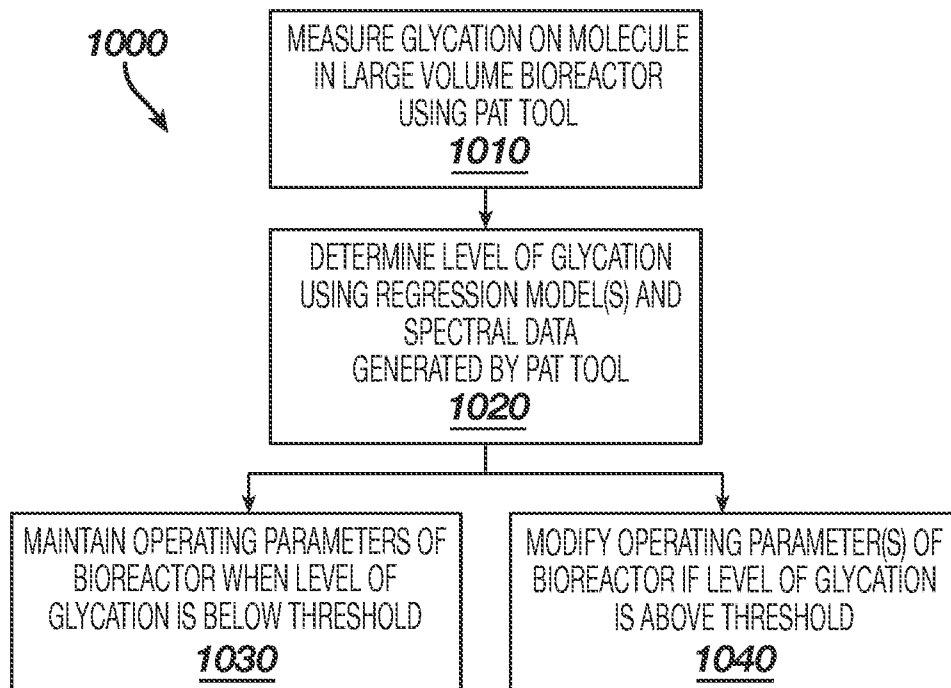
FIG. 5G

6/14

**FIG. 6**

**FIG. 7****FIG. 8**

8/14

**FIG. 9****FIG. 10**

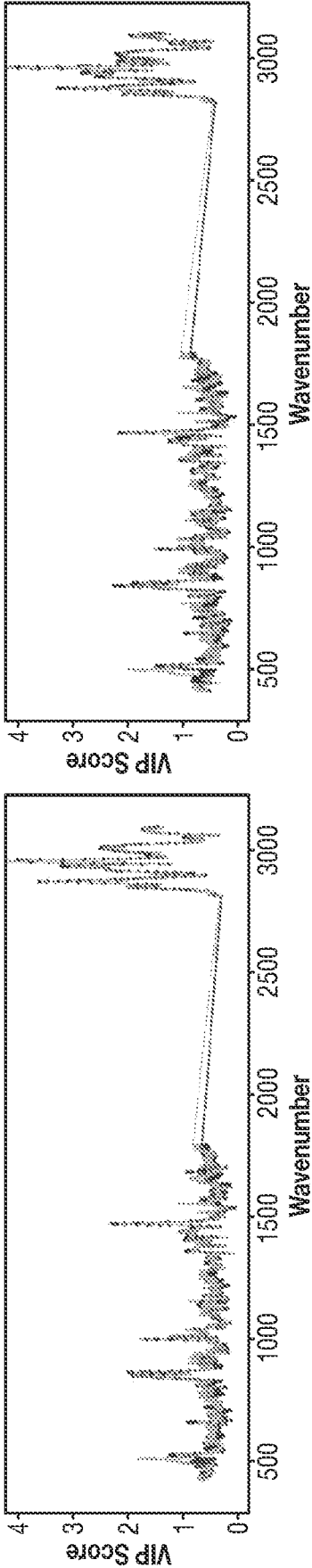
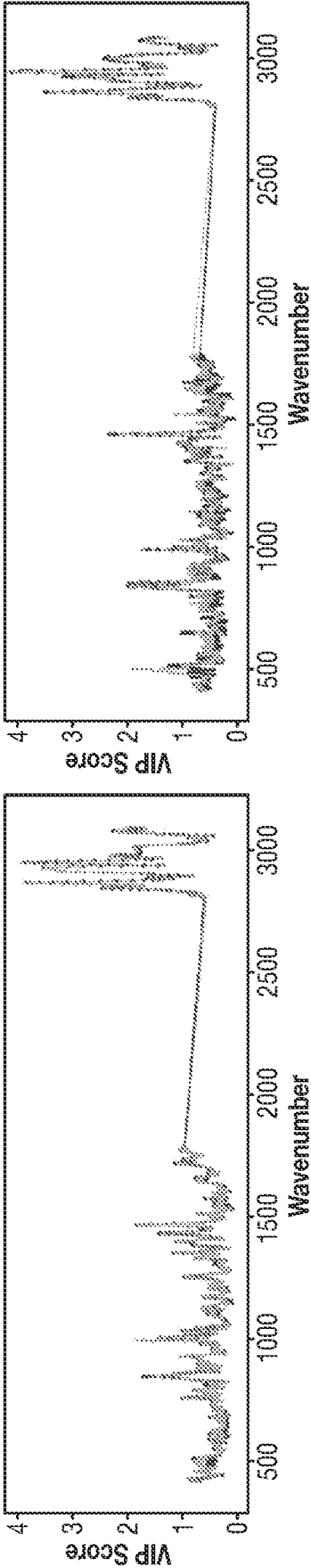
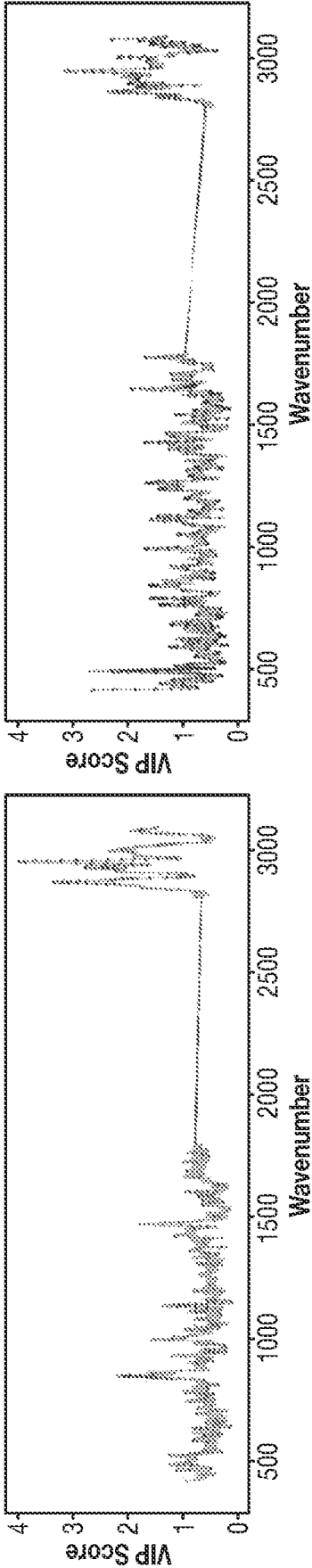


FIG. 11A

FIG. 11B

FIG. 11C

FIG. 11D

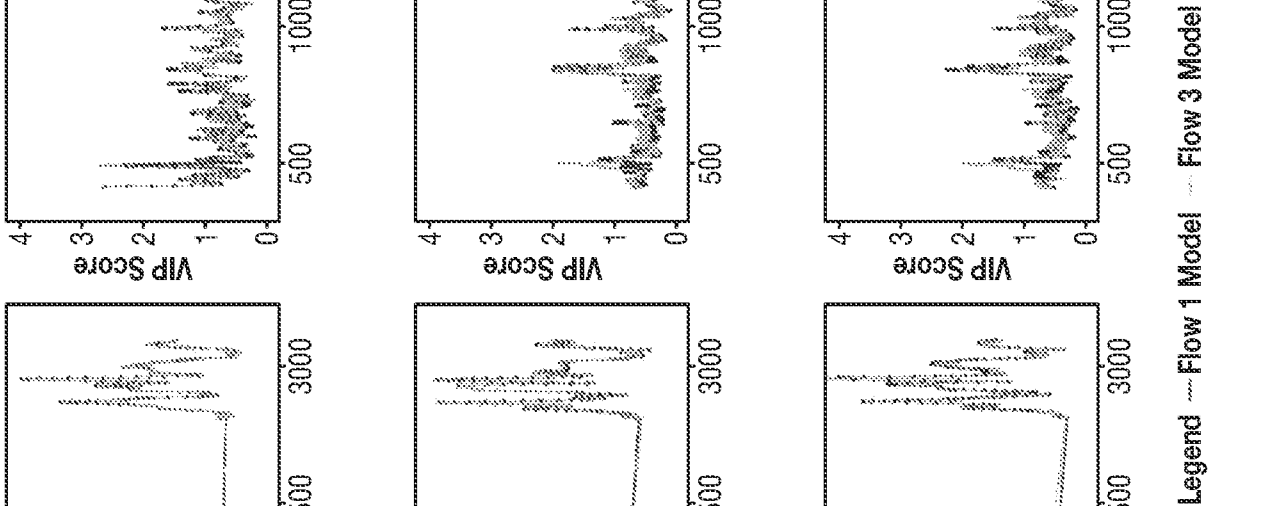


FIG. 11E

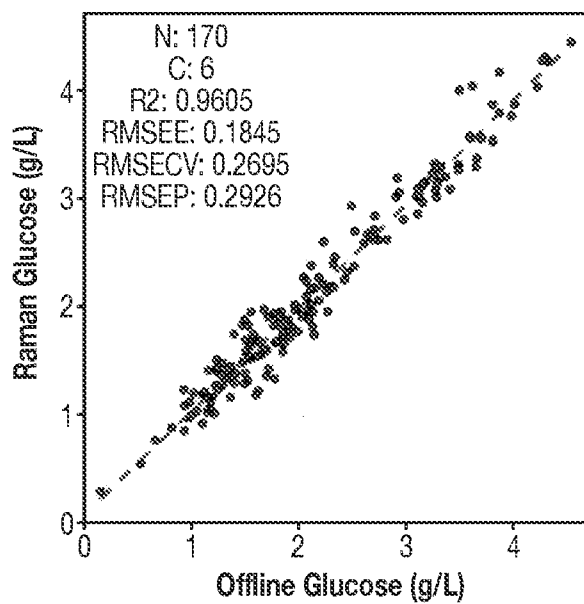
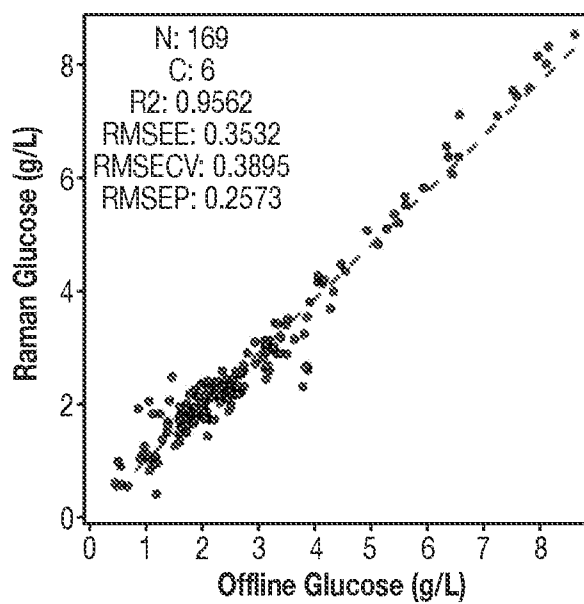
FIG. 11F

FIG. 11G

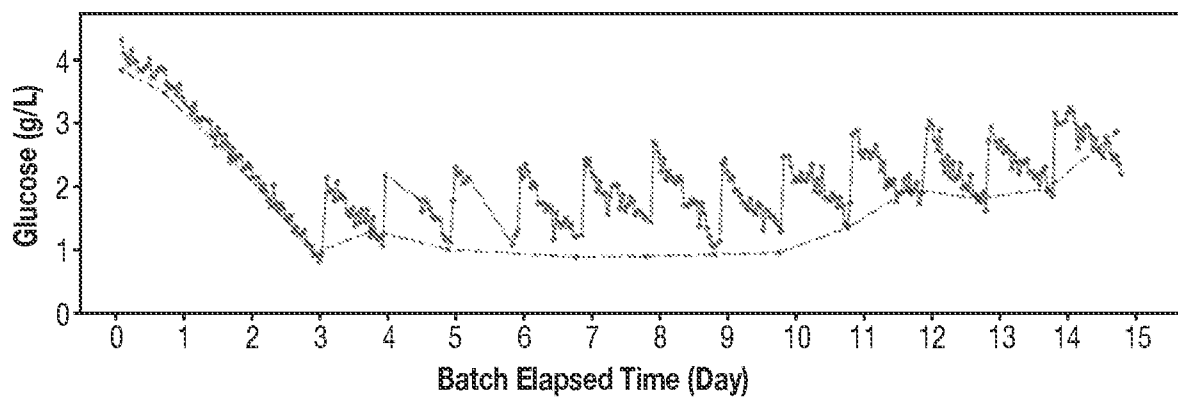
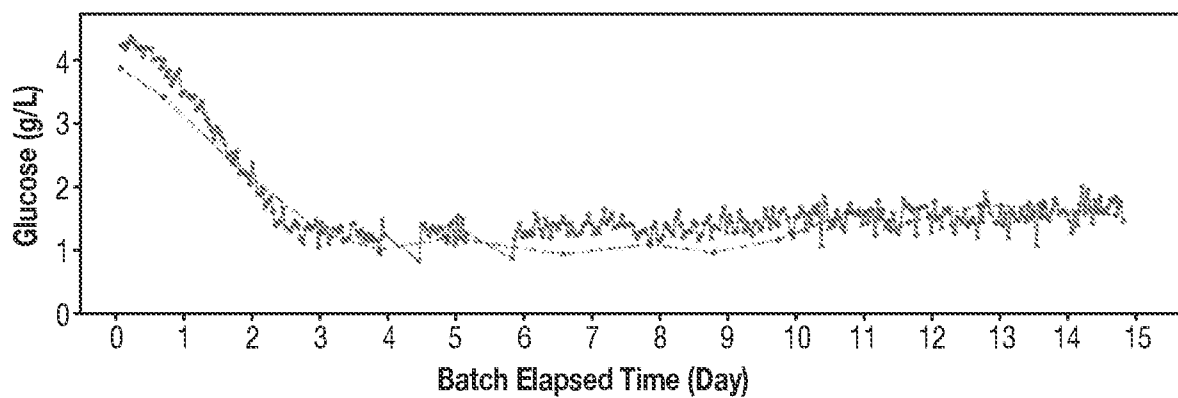
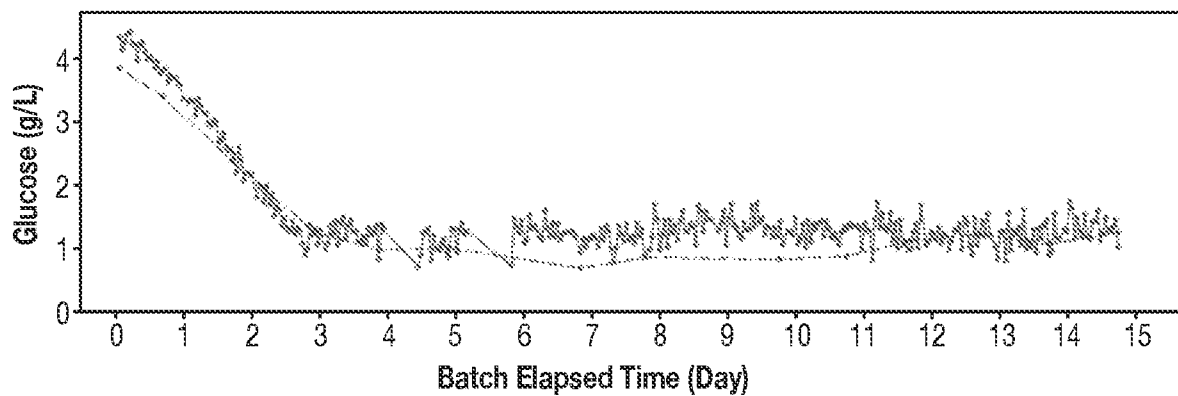
FIG. 11H

Legend — Flow 1 Model — Flow 3 Model

10/14

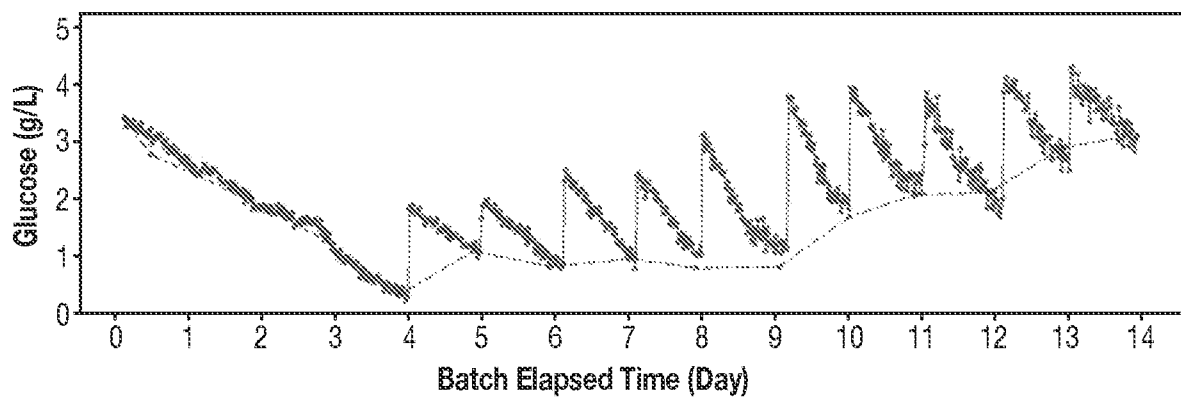
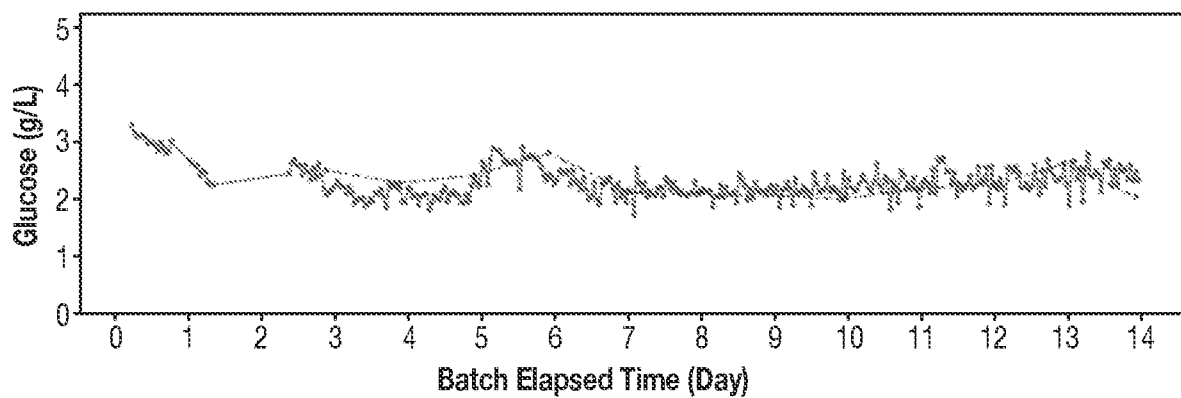
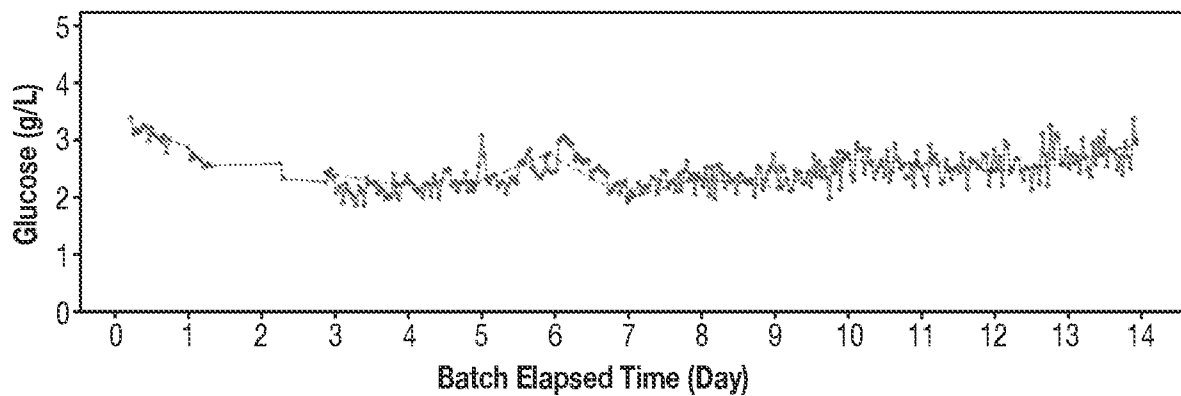
**FIG. 12A****FIG. 12B**

11/14

**FIG. 13A****FIG. 13B****FIG. 13C**

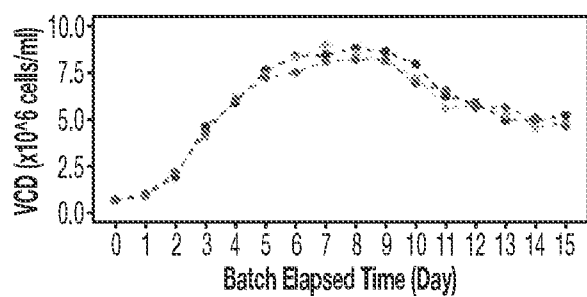
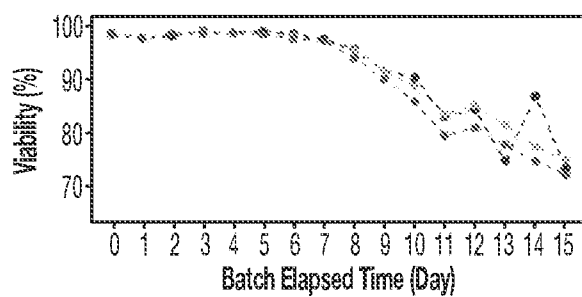
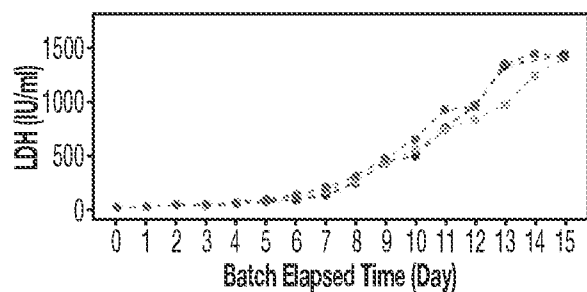
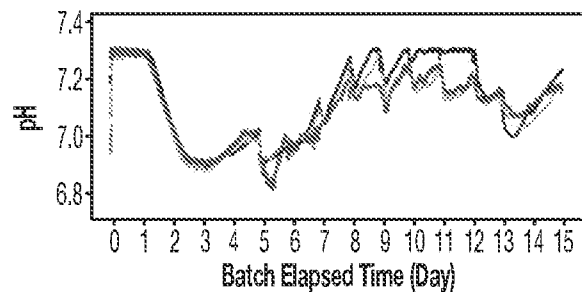
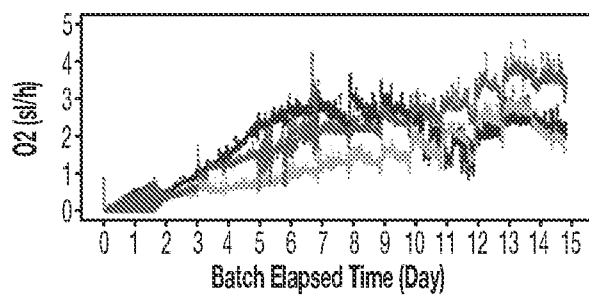
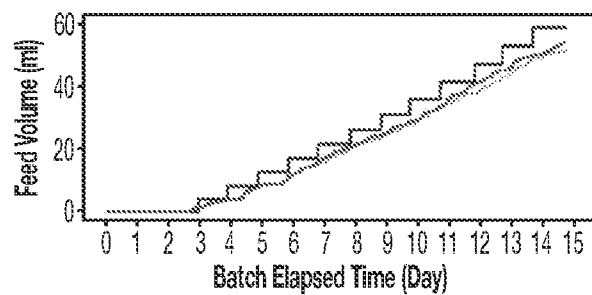
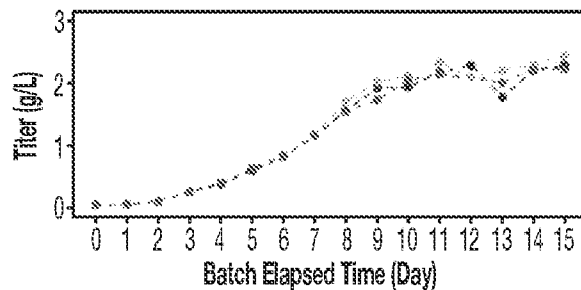
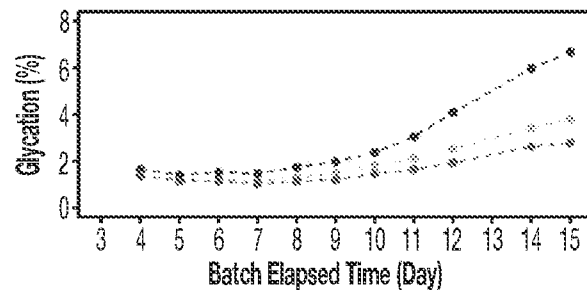
Legend ··· Offline Measurement — Raman Prediction

12/14

**FIG. 14A****FIG. 14B****FIG. 14C**

Legend --- Offline Measurement --- Raman Prediction

13/14

**FIG. 15A****FIG. 15B****FIG. 15C****FIG. 15D****FIG. 15E****FIG. 15F****FIG. 15G****FIG. 15H**

Legend $\diamond$ Batch A: Bolus Fed $\times$ Batch B: Glucose Control to 1g/L ∇ Batch C: Glucose Control to 1g/L

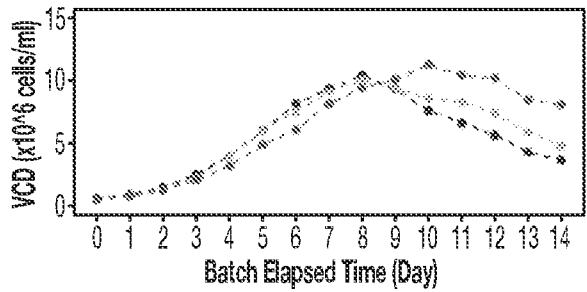


FIG. 16A

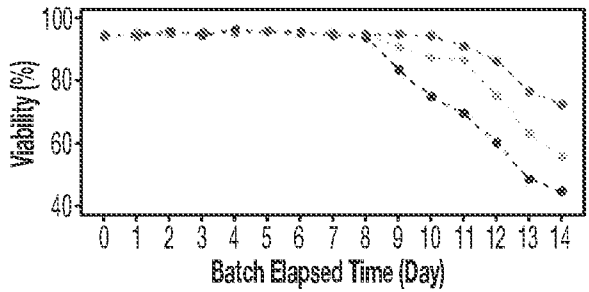


FIG. 16B

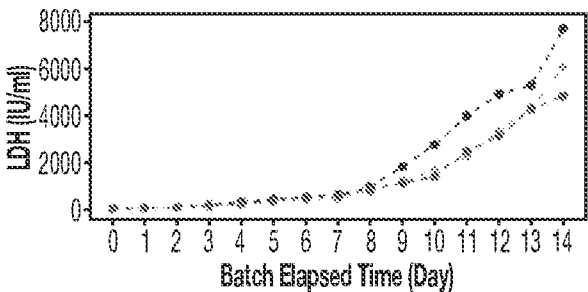


FIG. 16C

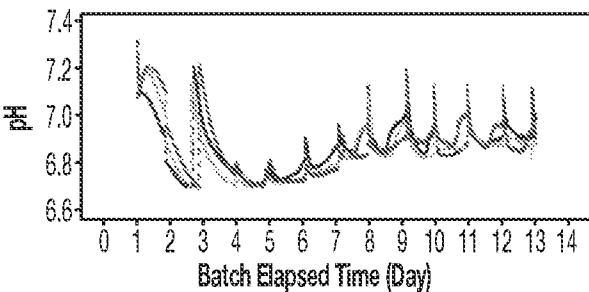


FIG. 16D

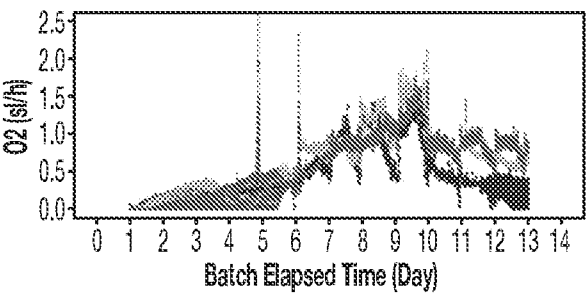


FIG. 16E

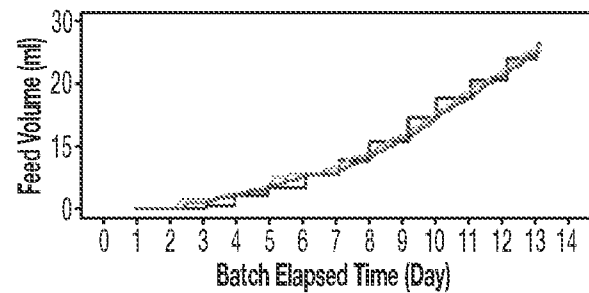


FIG. 16F

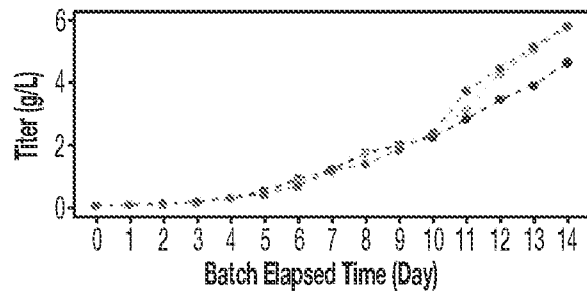


FIG. 16G

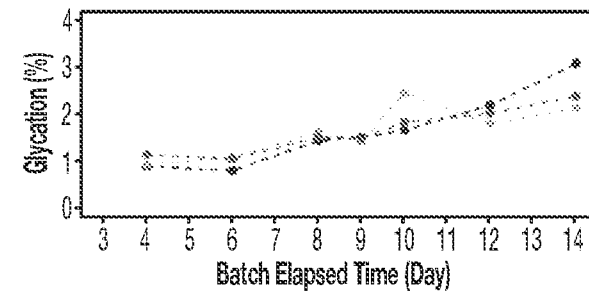


FIG. 16H

Legend ◊ Batch D: Bolus Fed ◊ Batch E: Glucose Control to 2g/L ◊ Batch F: Glucose Control to 2g/L

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 22/21289

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. G01N 33/68, G01N 21/65, G16B 15/00, G16C 20/20; ADD. C07K 14/00 (2022.01)

CPC - INV. G01N 33/68, G01N 21/65, G16B 15/00, G16C 20/20, C12P 21/005, G01N 2400/10, G01N 2440/38; ADD. C07K 14/00

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

See Search History document

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BREWSTER et al. "Monitoring the Glycosylation Status of Proteins Using Raman Spectroscopy" Anal. Chem. 2011, Vol. 83, No. 15, pp 6074-6081, especially, Abstract; page 6075, Col. 2, para 3; page 6076, Col. 2, para 3; page 6078, Col. 1, para 4; page 6078, Col. 2, para 1-2; page 6080, Col. 1, para 3-4	1-4
Y	US 2017/0355947 A9 (BIOGEN MA INC.) 14 December 2017 (14.12.2017) Abstract; para [0004]; para [0056]; para [0214]; para [0216]; claim 1; claim 9	1-4
A	US 2014/0349337 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 27 November 2014 (27.11.2014), whole document	1-4

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

10 June 2022

Date of mailing of the international search report

AUG 15 2022

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Kari Rodriguez

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 22/21289

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 5-15, 19-25, 29-37, 41-43
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I: Claims 1-4 directed to a method for determining a glycan structure on a glycosylated molecule.

Group II: Claims 16-18, 26-28 and 38-40 directed to a method or system of producing a glycosylated molecule.

The inventions listed as Groups I and II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Features

-----continued on supplemental sheet-----

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Box III: Observations where unity of invention is lacking:

Group I requires a method for determining a glycan structure on a glycosylated molecule, first bioreactors having a first volume equal to or below a first threshold, one or more second bioreactors having a second volume equal to or above a second threshold to result in measured spectral data; and determining, by at least one computing device using the generated one or more regression models and based on the measured spectral data, levels of the one or more glycan structures on the glycosylated molecule within one or more second bioreactors, not required by Group II.

Group II requires producing one or more glycosylated molecules comprising means for culturing a glycosylated molecule-producing cell line; means for measuring a level of one or more glycan structures and means for measuring a level of one or more glycan structures in a glycosylated molecule producing cell line, not required by Group I.

Common Technical Features

The features shared by the inventions listed as Group I and II are a glycan structure on a glycosylated molecule, obtaining, for each of a plurality of runs, levels for one or more glycan structures on the glycosylated molecule using a process analytical technology (PAT) tool, one or more regression models based on the obtained spectral data, determining, by at least one computing device.

However, these shared technical features do not represent a contribution over prior art, because the shared technical features are taught by an article entitled "Monitoring the Glycosylation Status of Proteins Using Raman Spectroscopy" by Brewster et al. (Anal. Chem. 2011, Vol. 83, No. 15, pp 6074-6081) (hereinafter 'Brewster').

Brewster discloses a glycan structure on a glycosylated molecule, obtaining, for each of a plurality of runs, levels for one or more glycan structures on the glycosylated molecule using a process analytical technology (PAT) tool, one or more regression models based on the obtained spectral data, determining, by at least one computing device (Abstract- "the characterization of these biopharmaceuticals poses a significant analytical challenge. In particular, monitoring posttranslational modifications (PTMs), such as glycosylation, is an important aspect of this characterization because these glycans can strongly affect the stability, immunogenicity, and pharmacokinetics of these biotherapeutic drugs. Raman spectroscopy is a powerful tool, with many emerging applications in the bioprocessing arena. Although the technique has a relatively rich history in protein science, only recently has Raman spectroscopy been investigated for assessing posttranslational modifications, including phosphorylation, acetylation, trimethylation, and ubiquitination. In this investigation, we develop for the first time Raman spectroscopy combined with multivariate data analyses, including principal components analysis and partial least-squares regression, for the determination of the glycosylation status of proteins"; page 6076, Col. 1, para 2; page 6076, Col. 2, para 1-3).

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Groups I and II therefore lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

Item 4 (continued)

Claims 5-15, 19-25, 29-37 and 41-43 are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).