On-demand manufacturing of clinical-quality biopharmaceuticals

Laura E Crowell[1,2,13], Amos E Lu[2,13], Kerry R Love[1,13], Alan Stockdale[1], Steven M Timmick[3–5], Di Wu[6], Yu (Annie) Wang[6], William Doherty[1], Alexandra Bonnyman[1], Nicholas Vecchiarello[3,4], Chaz Goodwine[3,4], Lisa Bradbury[7], Joseph R Brady[1,2], John J Clark[1,8], Noelle A Colant[1,2,9], Aleksandar Cvetkovic[7,10], Neil C Dalvie[1,2], Diana Liu[1], Yanjun Liu[6], Craig A Mascarenhas[1,11], Catherine B Matthews[1,2], Nicholas J Mozdzierz[1,2], Kartik A Shah[1], Shiaw-Lin Wu[12], William S Hancock[6], Richard D Braatz[2], Steven M Cramer[3,4] & J Christopher Love[1,2]

Conventional manufacturing of protein biopharmaceuticals in centralized, large-scale, single-product facilities is not well-suited to the agile production of drugs for small patient populations or individuals. Previous solutions for small-scale manufacturing are limited in both process reproducibility and product quality, owing to their complicated means of protein expression and purification[1–4]. We describe an automated, benchtop, multiproduct manufacturing system, called Integrated Scalable Cyto-Technology (InSCyT), for the end-to-end production of hundreds to thousands of doses of clinical-quality protein biologics in about 3 d. Unlike previous systems, InSCyT includes fully integrated modules for sustained production, efficient purification without the use of affinity tags, and formulation to a final dosage form of recombinant biopharmaceuticals. We demonstrate that InSCyT can accelerate process development from sequence to purified drug in 12 weeks. We used integrated design to produce human growth hormone, interferon α-2b and granulocyte colony-stimulating factor with highly similar processes on this system and show that their purity and potency are comparable to those of marketed reference products.

Biologic medicines, such as recombinantly expressed cytokines, hormones, replacement enzymes, blood factors or antibodies, are routinely used to treat cancer, autoimmune disorders and rare diseases. Increasingly, protein biologics are tailored to small groups of patients based on an understanding of the underlying biology of their disease[5]. The need for small numbers of doses of many products poses a challenge to conventional manufacturers, who produce drugs in large volumes to achieve economies of scale[6]. Furthermore, different classes of biopharmaceuticals (for example, enzymes, hormones, vaccines) generally require unique customized processes for each molecule from expression to purification, constraining commercial facilities to a single class of product.

New technologies to manufacture many different pharmaceutical-quality biologics in small quantities with efficiency and agility are needed to make precision biologic medicines both available and economically feasible[7]. Technologies such as automated laboratory-scale batch processes, *in vitro* transcription and translation, and microfluidics can rapidly produce limited quantities of different biomolecules on demand[1–4]. While some of the products generated show biological activity, they lack sufficient quality attributes for clinical use, including identity, purity, safety and potency as required by regulatory agencies. To address this need, we developed an automated multiproduct manufacturing system capable of rapidly producing clinical-grade recombinant proteins and requiring only minimal reconfiguration to make different biopharmaceuticals. Unlike previous solutions, InSCyT comprises fully integrated and automated modules for sustained production, efficient purification of the native protein, and final formulation for parenteral use. The yeast-based production module allows both rapid production of tens to hundreds of doses in under 80 h and sustained production for >100 h of up to thousands of doses as needed, whereas cell-free solutions[1–3] provide only short-term expression. The multistage purification module (and associated methods for rapid process development of suitable purification sequences) enables purification of multiple clinical-grade products using highly similar processes without requiring affinity tags that alter the drug sequence and present risks of immunogenicity. In contrast, previous approaches used affinity tags for purification or had no integrated purification step. Finally, the integrated formulation module prepares protein biopharmaceuticals in a final dosage form. These

[1]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [2]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [3]Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA. [4]Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York, USA. [5]GlaxoSmithKline, King of Prussia, Pennsylvania, USA. [6]Barnett Institute and Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts, USA. [7]Pall Life Sciences, Westborough, Massachusetts, USA. [8]Biogen, Cambridge, Massachusetts, USA. [9]Department of Biochemical Engineering, University College London, London, England. [10]Sanofi, Framingham, Massachusetts, USA. [11]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [12]BioAnalytix Inc., Cambridge, Massachusetts, USA. [13]These authors contributed equally to this work. Correspondence should be addressed to J.C.L. (clove@mit.edu).

Received 16 November 2017; accepted 27 August 2018; published online 1 October 2018; doi:10.1038/nbt.4262

features, combined with the use of integrated design principles to create simplified processes, facilitate the production of recombinant proteins with quality attributes sufficient for clinical use.

We selected *Pichia pastoris* as our expression host because it can grow quickly to high cell densities and efficiently secrete recombinant proteins[8]. Other advantages of *P. pastoris* include low levels of secreted host-cell proteins; little to no risk of viral contamination; validated expression of myriad proteins, including therapeutics approved by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA); and the capability for human-like post-translational modifications in engineered strains[9,10].

InSCyT uses fluidically connected modules for fermentation, multi-stage chromatography, and ultrafiltration and/or diafiltration, as well as integrated sensors and system controllers for system-wide programmed operations (**Fig. 1a,b** and **Supplementary Fig. 1**). We implemented continuous fermentation by perfusion to reduce the volume of the bio-reactor and enable high space–time yields[11]. To this end, we adapted a sub-liter benchtop bioreactor for in-tank perfusion and equipped it with sensors to control input and output flows, pH, temperature, impeller speed and dissolved oxygen (**Fig. 1c**). The bioreactor was con-nected to a module for inline pH adjustment of the cell culture fluid before chromatographic separations; this module allowed the balanc-ing of flow rates between those for production and purification. An integrated module for purification was designed to enable either two or three stages of chromatographic separation (**Fig. 1c**). This module allowed straight-through processing with no intermediate holding tanks or adjustments between purification steps. Our design simpli-fies the operation of the module relative to traditional purifications, wherein multiple intermediate procedures are often required to adjust pH, conductivity, concentration and composition of fluids between steps of purification. The final module in the system is a tangential flow-filtration system for buffer exchange and formulation to a final liq-uid dosage form of the product (**Fig. 1c**). A custom integrated software architecture unified operation of all three modules with appropriate controls as a fully automated single system.

We built and used three independent InSCyT systems to demon-strate consistent operational performance for both production and purification processes (**Fig. 1d,e**). We first produced and purified the common biologic medicine human growth hormone (hGH), used to treat growth deficiencies[12]. The extensive knowledge available about this drug's quality, safety and potency aided our assessment of InSCyT and the novel process we developed for production of hGH. Our production process used two straight-through stages of chromato-graphic purification and required fewer than half as many operations as the Innovator process (**Supplementary Fig. 2**). We implemented our process on three separate InSCyT systems. After inoculation, each process ran 'hands-free' and produced >100 doses of formulated product in less than 1 week, with initial yields for the process ranging from 27–31% on each system (**Fig. 2a**).

We compared the biophysical and biochemical attributes of our puri-fied product from multiple time points on each system to a marketed drug substance using multiple analytics commonly used to establish identity, potency, safety and purity of therapeutic proteins[13] (**Fig. 2b,c** and **Supplementary Fig. 3**). We confirmed the protein sequence (100% coverage) by mass spectrometry (**Supplementary Fig. 4**). Potency of InSCyT hGH was comparable to the National Institute for Biological Standards and Control World Health Organization (WHO) 98/574 ref-erence standard (96–104%) using a cell-based proliferation assay. Key contributors to product safety include the levels of potentially immuno-genic product-related impurities, such as aggregates, and the presence of process-related impurities, including host-cell proteins (HCPs) and

host-cell DNA. Minimal high-molecular-weight species were present in the InSCyT hGH product (<0.5%), and levels of process-related impurities were each below typical values for clinical-stage develop-ment (1,000 p.p.m. for HCPs and 10 ng/dose for DNA)[14–16]. Regulatory agencies usually consider limits for host-cell proteins on a case-by-case basis, although *in vitro* studies using peripheral blood mononuclear cells from both healthy and diseased individuals have shown that HCP levels up to 4,000 p.p.m. from Chinese hamster ovary (CHO) cells do not pose a higher immunogenicity risk than a highly purified mono-clonal antibody (<50 p.p.m.)[15]. A small percentage (average: 6%) of a proteolytically cleaved form of hGH was also observed (two-chain variant); this natural form is both highly potent and previously has been determined as clinically irrelevant[17]. Further assessment for product-specific impurities by tandem liquid chromatography/mass spectrom-etry (LC–MS), however, showed that InSCyT hGH was not comparable to marketed products because of increased levels of deamidation and oxidation (**Fig. 2c** and **Supplementary Fig. 5**).

Given the consistency of both the operation of individual InSCyT systems and the hGH products produced by each system, we attrib-uted the deamidation and oxidation observed in our product to process parameters, rather than the InSCyT system specifically. We therefore modified our process for hGH without any significant hardware changes by changing to a defined cultivation medium, adjusting the set point for dissolved oxygen and eliminating agitation in the surge tank. Using this adjusted process, a single InSCyT system produced nearly 50 maximum weight-based doses of hGH in 75 h (**Supplementary Fig. 6**). Oxidation and deamidation were reduced to below 1% at each residue, while all other quality attributes were maintained (**Fig. 2c,d** and **Supplementary Fig. 6**). Process yield also increased to nearly 80%. Together, these data demonstrate the capability of our manufacturing system to rapidly and reproducibly produce tens of doses of a potent and pure form of a bio-logic drug in an automated, short production cycle.

We next sought to demonstrate on-demand production of hGH. The upstream perfusion process was operated fully automated for 240 h. In addition to the initial purification and formulation cycle on day 3 (described above), two more cycles of purification and formulation were performed on demand during days 6 and 10 (**Supplementary Fig. 6**). Each on-demand cycle produced between 50 and 75 doses of hGH within 12 h, with product quality and yields similar to those of the batch pro-duced on day 3 (**Fig. 2c,d** and **Supplementary Fig. 6**). (We attributed the reduced yield (50%) and increased host-cell protein levels observed dur-ing the day 6 cycle to overloading the capture column, which we adjusted before the day 10 cycle.) Further optimization of the process and column sizing could improve the consistency from batch to batch. Nonetheless, the ability to produce small lots of this product on demand shows the potential for manufacturing medicines as needed and highlights the sta-bility of yeast-based bioprocesses in continuous operations.

Biopharmaceuticals typically require custom manufacturing proc-esses that vary widely, especially for proteins other than monoclonal antibodies, and require unique facility designs[11]. This constraint limits facility flexibility to provide additional products, which would be essential for on-demand supply. We therefore assessed whether our modular manufacturing system and choice of host could readily produce other molecules with no substantial hardware alterations. We selected interferon-α (IFN-α)-2b as a second example. This potent 19.2-kDa cytokine is used in both monotherapies and drug combina-tions to treat cancer and hepatitis, and is produced commercially in a specialized 13-step process using *E. coli*[18] (**Supplementary Fig. 2**). Owing to the ease of targeted transgene insertion and simplicity of upstream process development in *P. pastoris,* we developed a draft fer-mentation process for producing secreted IFNα-2b less than 4 weeks
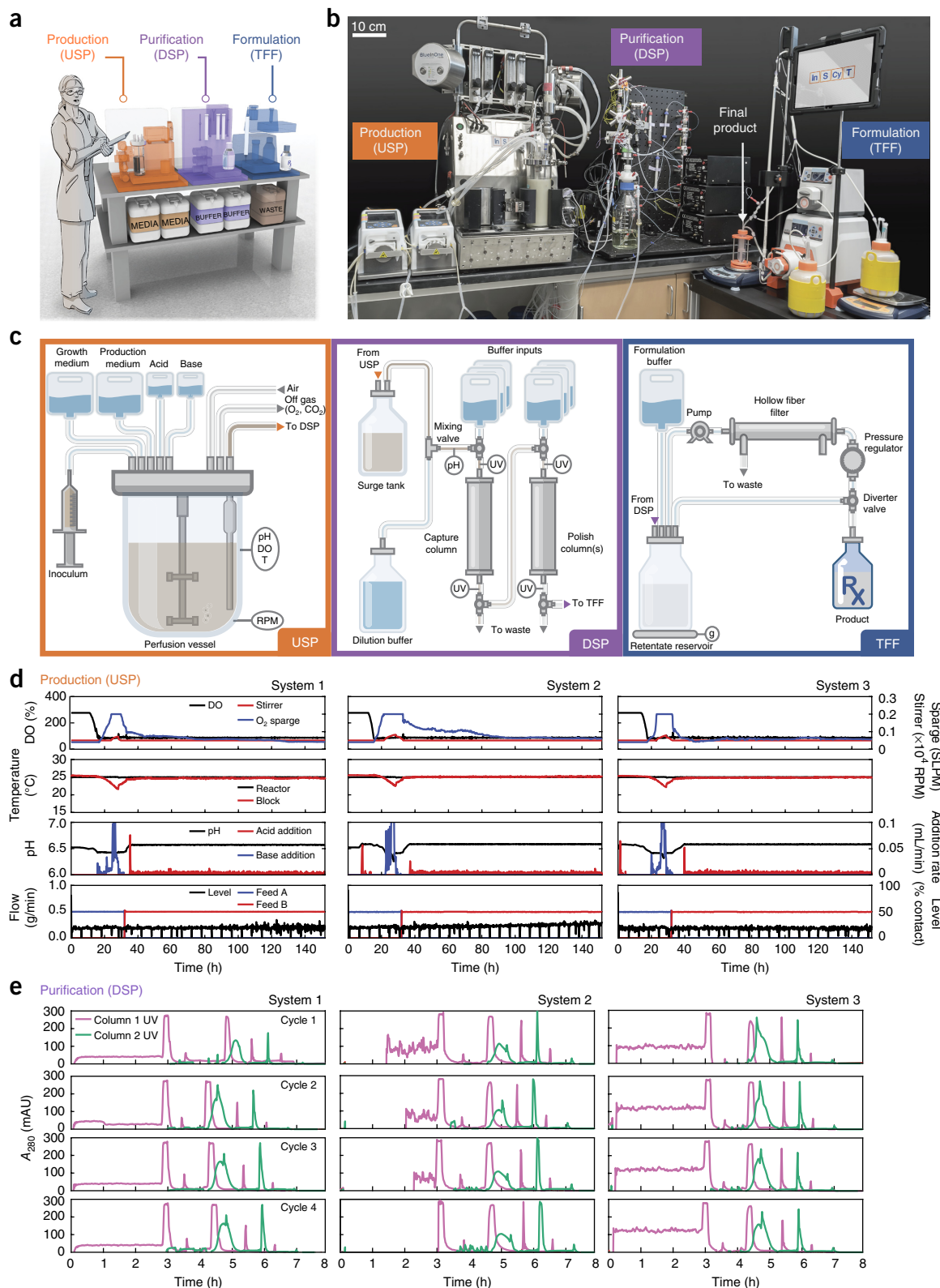
Figure 1 Schematic of the InSCyT system for on-demand biomanufacturing and demonstration of consistent operation across three distinct InSCyT systems. (**a**) To-scale rendering of the InSCyT system. Human figure is approximately 5 feet 7 inches (170 cm) tall. (**b**) Photograph of an operational InSCyT system. (**c**) Detailed schematic of the InSCyT system, including interactions between modules and key control points for the production (upstream processing, USP), purification (downstream processing, DSP) and formulation (tangential flow filtration, TFF) modules. DO, dissolved oxygen; T, temperature. (**d**,**e**) Process parameter profiles collected by the control software from the production (USP) module (**d**) and the purification (DSP) module (**e**) of three separate InSCyT systems during hGH fermentation. SLPM, standard liters per minute.

after identifying the product sequence (**Fig. 3a**). We have found that *P. pastoris* routinely secretes a consistent set of host-cell proteins along with the heterologous product during fermentation to yield a high level of initially pure product (>80%)[19]. This feature made it possible to develop an *in silico* tool to predict draft multistage purification processes[20]. With this *in silico* tool, we selected a process to purify IFNα-2b within another 4 weeks. The procedures developed for both production and purification did not require modifications to the InSCyT system itself.

Initial biophysical analyses of IFNα-2b produced in our first run on the InSCyT system indicated minimal high molecular weight species (0.34%) and process-related impurities below typical values for clinical development (**Fig. 3c**). A cell-based viral replication assay demonstrated that the potency of our IFNα-2b was the same as or greater than that of a reference drug substance. InSCyT-generated IFNα-2b was highly potent (134%) in part owing to a naturally occurring C-terminal truncation known to increase potency[21]. Assessment for purity by LC–MS and reversed-phase liquid chromatography (RPLC), however, showed that our product quality was not sufficient due to the presence of oxidized forms that could potentially promote aggregation and immunogenicity[22] (**Fig. 3c**). These data show that rapid production of biomolecules with acceptable bioactivity is necessary, but not sufficient, to define a clinical-quality biologic product.

To address these attributes, we optimized process conditions on InSCyT, performing experiments on individual modules simultaneously with fully integrated experiments comprising connected modules (**Fig. 3a** and **Supplementary Fig. 7**). After 27 d of process development, a final run showed that oxidation was reduced to <1.5% at all residues (**Fig. 3b,c**). During this run the system produced nearly 8,000 formulated doses of IFNα-2b in less than 1 week (**Supplementary Fig. 8**). The identity and purity of the product was confirmed by multiple analytical methods at four time points during the campaign (**Fig. 3c** and **Supplementary Fig. 8**). RPLC, size exclusion chromatography and LC–MS showed levels of purity within targeted specifications; other analytics showed the specifications for safety and potency of our product were also achieved. Differences in chromatographic behavior of our product were confirmed by matrix-assisted laser desorption and ionization mass spectrometry as the naturally occurring C-terminal truncation mentioned previously[21] (**Supplementary Fig. 8**). The low overall process yield (~11%) is attributed to removal of an N-terminal product variant, which arises as a result of incomplete cleavage of the secretion leader sequence during expression by our host. Further engineering of the expression vector with alternative signal sequences could alleviate the expression of this variant[23].

Next we produced granulocyte colony stimulating factor (G-CSF), used to stimulate blood cell proliferation and reduce infections in cancer patients treated with myelosuppressive chemotherapy[24]. This drug has manufacturing challenges due to complex folding and a propensity for aggregation and oxidation at specific amino acid residues[25,26]. We designed a new process for production of G-CSF and implemented it on three InSCyT systems (**Supplementary Fig. 2**). After inoculation, each automated process yielded more than 165 doses of formulated drug in less than 100 h (**Fig. 4a** and **Supplementary Fig. 9**). Typical process yields for each cycle ranged from 70 to 90% (average 77%). We assessed biophysical and biochemical attributes of the product using multiple analytics to establish its identity, purity, safety and potency (**Fig. 4a** and **Supplementary Fig. 9**). InSCyT G-CSF was comparable to a drug substance from a licensed product. We confirmed the protein sequence (100% coverage) by mass spectrometry (**Supplementary Fig. 10**). Minimal high-molecular-weight species were present in the product (0.33–0.65%). Levels of process-related contaminants in our formulated G-CSF were each below values typical for early-stage clinical development (1,000 p.p.m. for HCPs and 100 pg/dose for DNA)[14–16]. Potency of the InSCyT G-CSF was comparable to that of the NIBSC WHO 09/136 reference standard (89.6–141.1%) in a cell-based proliferation assay. Our product contained a minor variant comprising an N-terminal truncation and was a mixture of aglycosylated and glycosylated forms (**Supplementary Fig. 9**). Neither of these variants are likely to be clinically meaningful, as both truncations and glycosylation have been observed in licensed products without impact on product activity or safety[27,28]. Overall, these data demonstrate that InSCyT can rapidly and consistently produce therapeutic proteins that are comparable to currently marketed products.
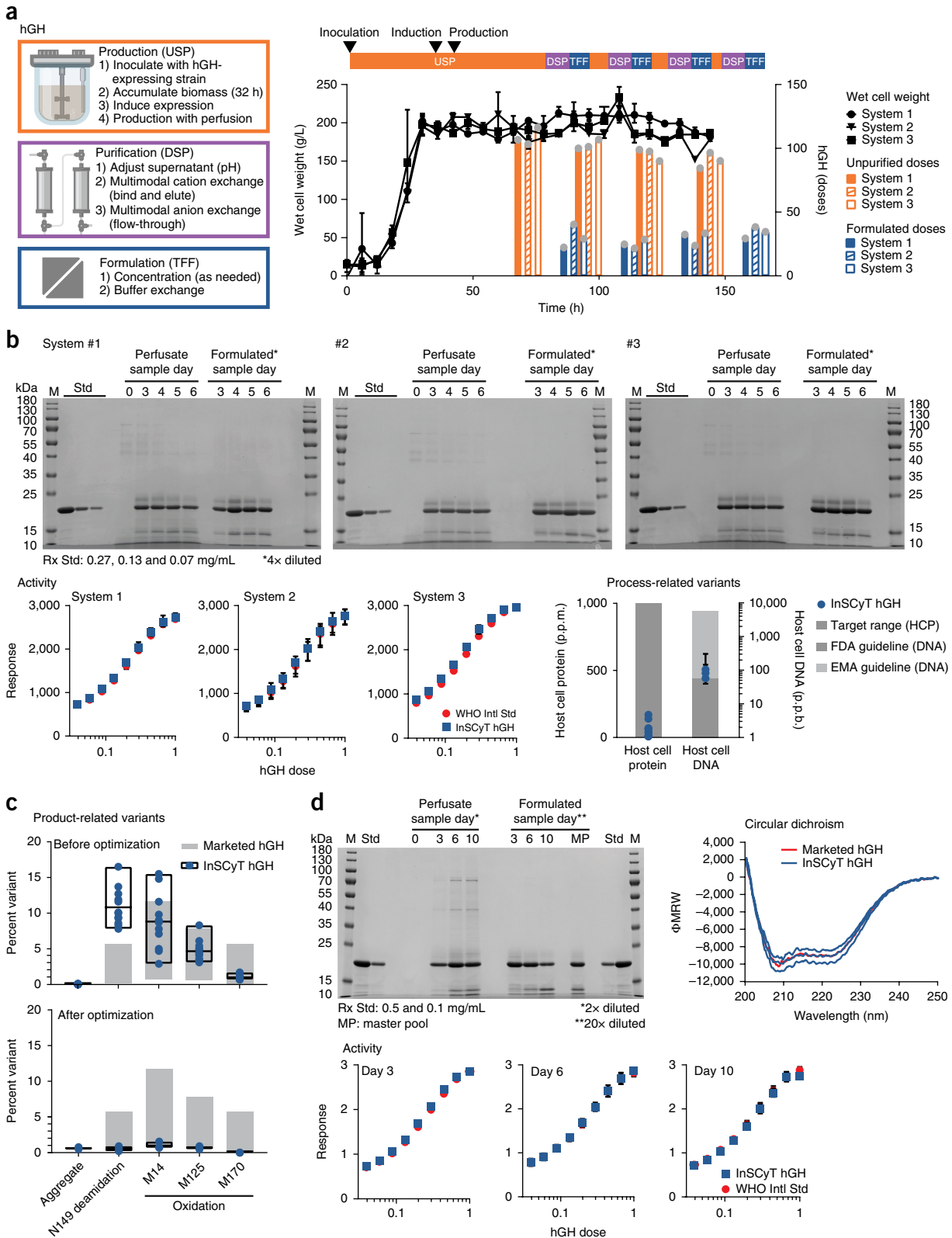
We performed further nonclinical studies with InSCyT-produced G-CSF to provide a framework for future clinical development. We assessed the pharmacokinetics, pharmacodynamics and toxicology of the InSCyT-produced G-CSF by comparing our product to a licensed product (Neupogen) in a rat model. We found InSCyT-produced G-CSF to be comparable to Neupogen in neutrophil activation during a single-dose administration study (**Fig. 4b**). InSCyT G-CSF and Neupogen showed no statistically significant difference in pharmacokinetic profile when administered at the same dose (Kolmogorov–Smirnov test, $P = 0.9963$) (**Fig. 4b**). A 5-d repeat-dosing study also showed our product was comparable to Neupogen in toxicity as based on survival, clinical signs, body weight, quantitative food consumption, hematology, serum chemistry, organ weights and macroscopic findings (**Fig. 4b** and **Supplementary Fig. 11**). In all these studies, no abnormal signs of toxicity, including injection site inflammation, were observed in any animals dosed with InSCyT G-CSF

Figure 2 Production of hGH on the InSCyT system. Dose size used was 1.75 mg (ref. 12). Center values and error bars represent the mean and range, respectively, of technical triplicates unless otherwise indicated. (**a**) Process flow chart (left) and timeline and yields (right) for production of hGH using InSCyT. Wet cell weight (black), unpurified (orange) and formulated (blue) doses of hGH produced are shown. Gray circles represent individual data points. (**b**) Product quality analyses for InSCyT-produced hGH before optimization, alongside a reference drug substance from a licensed hGH product produced in *E. coli*. SDS–PAGE (12% Tris-glycine) analysis of samples from the USP during biomass accumulation and production (perfusate samples), final formulated samples, and the reference (Std); M, molecular mass marker. Activity of InSCyT hGH alongside the WHO international standard (NIBSC 98/574). The final formulated sample (day 6) was analyzed from each system. Quantification of host-cell protein (HCP) and host-cell DNA impurities in formulated InSCyT hGH. Host-cell protein limits are shown as a target range[14,15]. Host-cell DNA guidelines are based on 100 pg/dose (FDA) and 10 ng/dose (EMA)[16,30]. For host-cell protein data, each point represents a unique sample (12 points total: 4 time points from each of three InSCyT systems). For host-cell DNA, each point represents a single pooled sample from each system comprising equal volumes of samples from each time point (3 points total, 1 per system). (**c**) Analysis of product-related variants in formulated InSCyT hGH before (top) and after optimization (bottom) alongside levels typically found in marketed products (**Supplementary Fig. 5**). Each data point represents a unique sample; there are 12 data points for runs before optimization (four time points from each of three InSCyT systems) and 3 data points for runs after optimization (three time points from a single InSCyT system). Black boxes represent the range of InSCyT hGH samples, with a line at the mean. (**d**) Product quality analyses for InSCyT-produced hGH after optimization alongside reference drug substance from a licensed hGH product. SDS–PAGE (12% Tris-glycine) analysis of samples from the USP during biomass accumulation and production (perfusate samples), final formulated samples, and the reference (Std). Activity of InSCyT hGH alongside the WHO international standard (NIBSC 98/574). Secondary structure analysis of InSCyT hGH (individual formulated samples from days 3, 6 and 10) and the reference hGH standard using circular dichroism; MRW, mean residue ellipticity.

(58 in total). Together, these data suggest that the InSCyT-produced G-CSF has potency *in vivo* and that potentially immunogenic process-related impurities are appropriately minimized.

InSCyT can produce a variety of clinical-quality recombinant therapeutic proteins in a liquid dosage form through integrated production, purification and formulation under a single control architecture. The efficient secretion of proteins by *P. pastoris*, combined with a holistic design of purification sequences, enabled processes for hGH, IFNα-2b and G-CSF that reduced the total number of processing steps by 45% or more and did not require refolding, excursions in pH or other
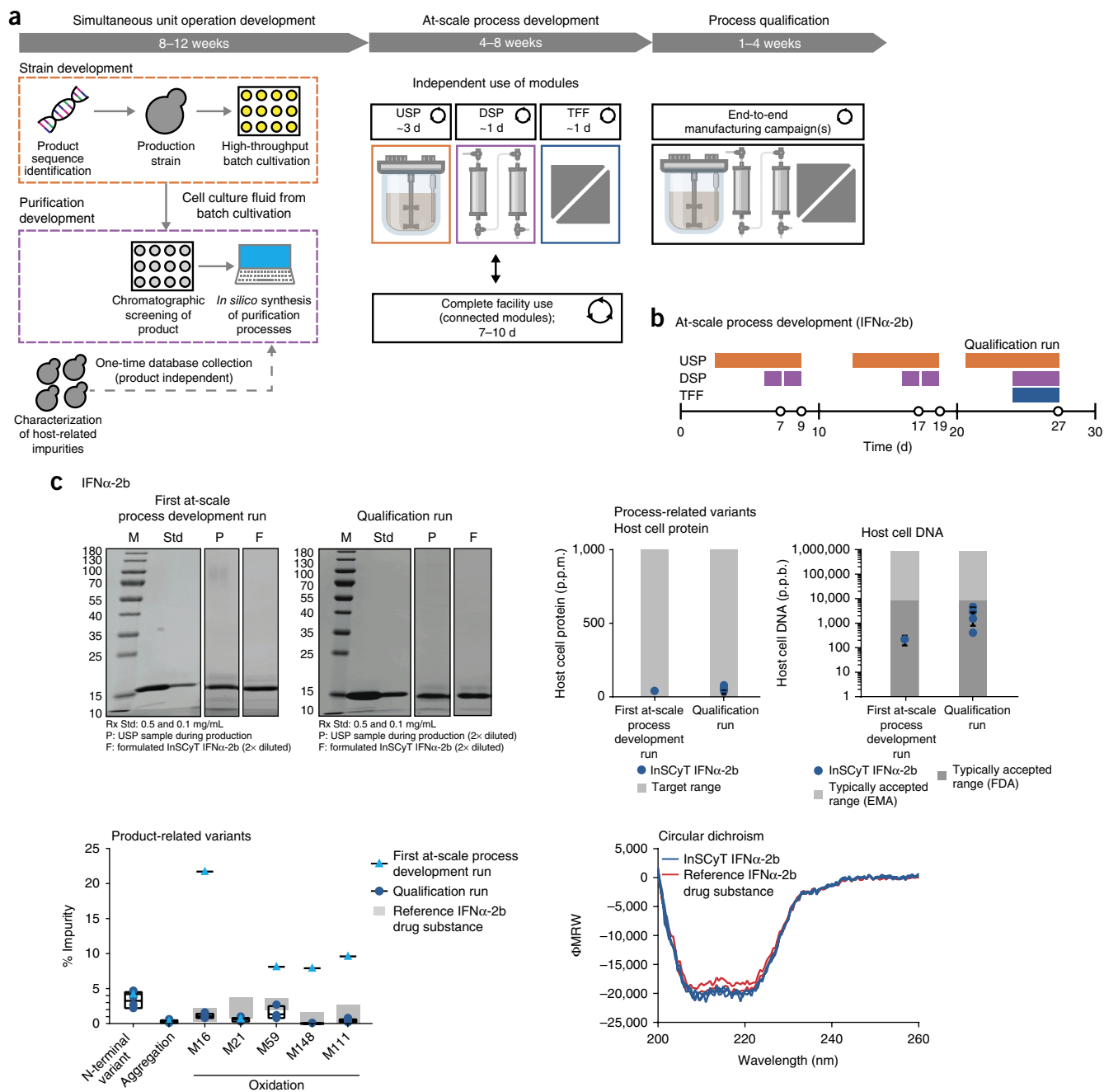

a hGH

Production (USP)
1) Inoculate with hGH-expressing strain
2) Accumulate biomass (32 h)
3) Induce expression
4) Production with perfusion

Purification (DSP)
1) Adjust supernatant (pH)
2) Multimodal cation exchange (bind and elute)
3) Multimodal anion exchange (flow-through)

Formulation (TFF)
1) Concentration (as needed)
2) Buffer exchange

b

Rx Std: 0.27, 0.13 and 0.07 mg/mL   *4× diluted

Activity — System 1, System 2, System 3
● WHO Intl Std   ■ InSCyT hGH

Process-related variants
● InSCyT hGH
Target range (HCP)
FDA guideline (DNA)
EMA guideline (DNA)

c Product-related variants

Before optimization — Marketed hGH, InSCyT hGH

After optimization

Aggregate, N149 deamidation, M14, M125, M170 (Oxidation)

d

Rx Std: 0.5 and 0.1 mg/mL   *2× diluted
MP: master pool   **20× diluted

Circular dichroism
Marketed hGH
InSCyT hGH

Activity — Day 3, Day 6, Day 10
■ InSCyT hGH
● WHO Intl Std

Figure 3 Accelerated process development using the InSCyT system and production of IFNα-2b. Dose size was 12 µg (ref. 18). Center values and error bars represent the mean and range, respectively, of technical triplicates unless otherwise noted. (**a**) Process development timeline for new manufacturing processes using the InSCyT system, including simultaneous unit operation development (comprising strain development and purification development), at-scale process development (comprising simultaneous experiments on individual modules and on the fully integrated system) and process qualification. (**b**) Timeline for at-scale process development for IFNα-2b. Horizontal colored bars represent the modules that were used in each experiment (USP, orange; DSP, purple; TFF, blue). Each new bar represents a new set of experimental conditions on that module. (**c**) Product quality for InSCyT-produced IFNα-2b from the first at-scale run after initial unit operation development (first at-scale process development run) and the final qualification run alongside a reference drug substance produced in *E. coli*. SDS–PAGE (12% Tris-glycine) analysis of samples from the USP during production (P), a final formulated sample (F) and a reference drug substance (Std); M, molecular mass marker. Analysis of process-related variants in formulated InSCyT IFNα-2b (per **Fig. 2b**). Each data point represents a unique sample, including one data point from the first at-scale process development run and four data points from the qualification run (four time points from a single InSCyT system). Product-related variants detected in formulated InSCyT IFNα-2b alongside levels typically found in a reference drug substance (**Supplementary Fig. 5**). Black boxes represent the range of InSCyT IFNα-2b samples, with a line at the mean. Secondary structure analysis of InSCyT IFNα-2b (triplicate analyses of an individual sample from the qualification run) and reference drug substance (duplicate analyses of an individual sample) using circular dichroism; MRW, mean residue ellipticity.
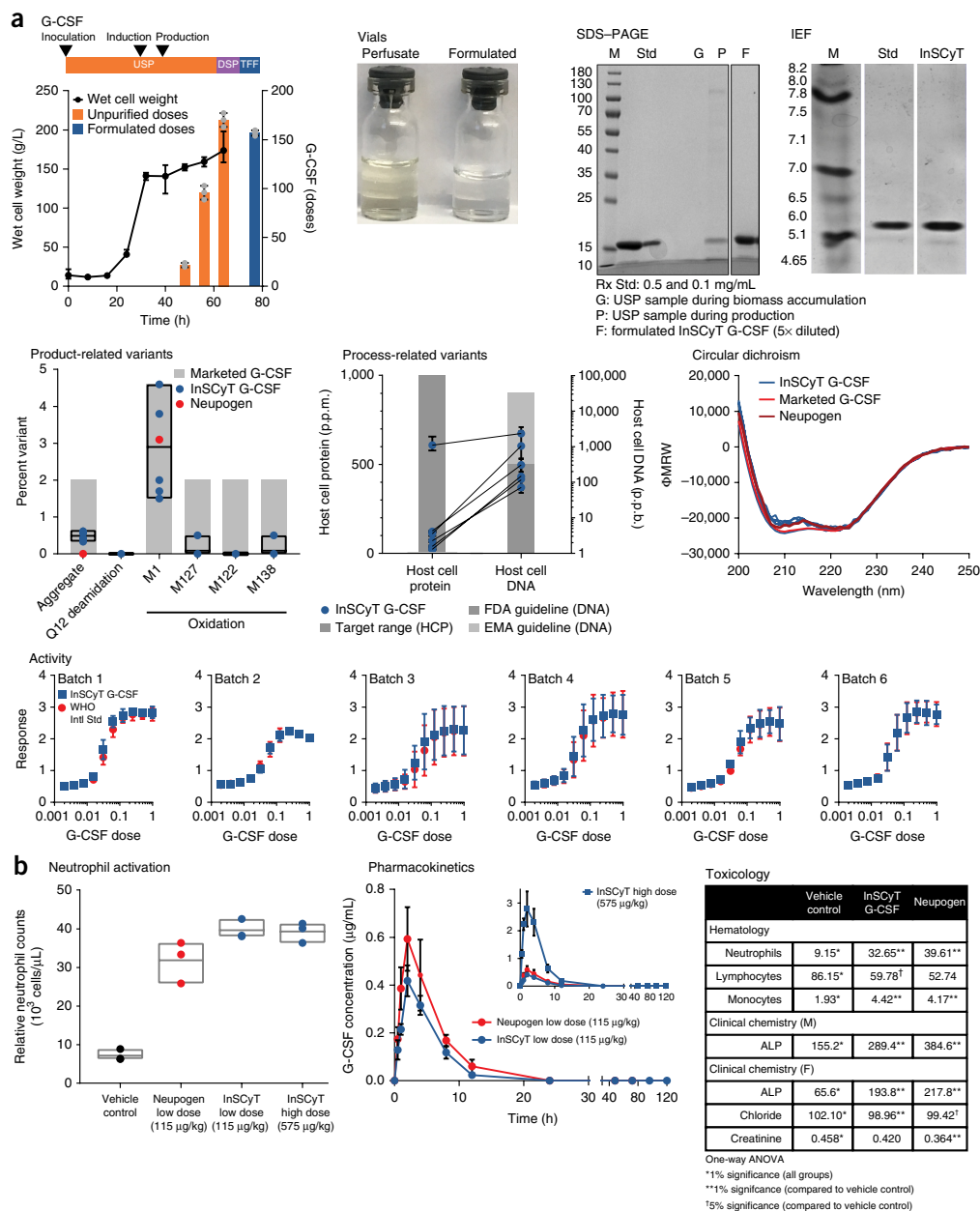
Figure 4 Production of G-CSF on three identical InSCyT systems. Dose size 300 µg (ref. 24). Center values and error bars represent the mean and range, respectively, of technical triplicates unless otherwise noted. (**a**) Timeline and yields for production of G-CSF using the InSCyT system for a single representative sample (batch 1). Wet cell weight (black circles) and cumulative unpurified (orange) and formulated (blue) doses of G-CSF are shown. Gray circles represent individual data points. Product quality for InSCyT-produced G-CSF alongside drug substance from a licensed product produced in *E. coli* and Neupogen (produced by Amgen in *E. coli*). A photograph of vials comparing material sampled from the USP (perfusate) to final formulated material (formulated). SDS–PAGE (12% Tris-glycine) analysis of batch 1 from the USP during biomass accumulation (G) and production (P), and a final formulated InSCyT sample (F) alongside drug substance from a licensed product (Std); M, molecular mass marker. Analysis of product purity by isoelectric focusing (IEF) for formulated batch 1. Gel analyses of batch 1 are representative of all six batches (**Supplementary Fig. 9**). Analysis of product-related variants and process-related variants. Each data point represents a unique batch (two time points from each of three distinct systems). Paired data points indicate analyses from a single batch. Product-related variants are shown alongside levels typically found in marketed products (**Supplementary Fig. 5**). Black boxes represent the range of InSCyT G-CSF samples, with a line at the mean. Process-related variants are shown alongside common guidelines (per **Fig. 2b**). Analysis of the secondary structures of InSCyT G-CSF (batches 1–6) and a reference drug substance from a licensed product using circular dichroism. Activity of InSCyT G-CSF alongside that of the WHO international standard (NIBSC 09/136). (**b**) Analysis of pharmacokinetics, pharmacodynamics and toxicology of InSCyT-produced G-CSF and a licensed product (Neupogen) in a rat model. Neutrophil activation and pharmacokinetic profile of low dose (115 µg/kg, $n = 3$ animals, $t_{1/2} = 2.1$ h) and high dose (575 µg/kg, $n = 3$ animals, $t_{1/2} = 4.6$ h) InSCyT G-CSF in rats compared to Neupogen (115 µg/kg, $n = 3$ animals, $t_{1/2} = 1.4$ h) (pharmacokinetics: $P = 0.9963$, Kolmogorov–Smirnov test). For neutrophil activation, gray boxes represent the range of three individual animals, with a line at the mean. For pharmacokinetics, center points and error bars represent the mean and range, respectively, of three individual animals. Summary of statistically significant results comparing the toxicology of InSCyT G-CSF and Neupogen to a vehicle control. Values represent the mean; s.d. and sample size can be found in **Supplementary Figure 11**. Statistical significance was determined by one-way ANOVA. ALP, alkaline phosphatase; MRW, mean residue ellipticity; M, male; F, female.

substantial changes to the protein itself during processing (**Supplementary Fig. 2**). We demonstrated fast cycles of process development to reach clinically relevant target specifications in 12 weeks, aided by testing production at scale in a modular and integrated manner on InSCyT. The combination of the manufacturing system with the demonstrated strategy for process development could facilitate the rapid transition of lead molecules into the clinic for translational studies and reduce subsequent iterations in process development and technology transfer for late-stage and commercial manufacturing.

Further engineering of InSCyT to comply with current good-manufacturing practices and concurrent development of an appropriate control strategy would enable its use for chemistry, manufacturing and controls of new drugs. A fill and finish module would enable product vial dispensing for simple administration to patients. Several relevant solutions have emerged, including systems from MedInstill and Vanrx. Modular facilities for housing manufacturing equipment, such as G-Con PODs and Germfree BioGO Modules, also are becoming widely available for aseptic containment of small-scale manufacturing facilities.

InSCyT could be used in its current form to produce many other products, such as monoclonal antibodies, vaccine components, nanobodies and other antibody-like proteins (for example, bispecific T-cell engagers, Fabs), blood products (such as erythropoietin) and therapeutic enzymes (for example, β-glucocerebrosidase). Other products, such as insulin or modified products such as antibody–drug conjugates or PEGylated versions of products, would require additional modules for enzymatic processing, chemical ligation or crystallization; such systems could include *de novo* synthesis of the key starting materials or active pharmaceutical ingredient as well[29]. Further integration of multiple units may also facilitate blended products of multicomponent vaccines or unique drug combinations tailored for applications for regional use or precision medicine.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

L.E.C., A.E.L. and K.R.L. designed experiments, analyzed data and wrote the manuscript. J.C.L., K.R.L., R.D.B., A.S., A.E.L., N.J.M. and L.E.C. designed and built the InSCyT system. A.E.L. and R.D.B. developed controls for the InSCyT system. L.E.C., W.D. and A.B. performed experiments on the InSCyT system and performed quality assessments. D.W., Y.A.W., Y.L., S.-L.W. and W.S.H. assessed quality by mass spectrometry and isoelectric focusing. S.M.T., N.V., C.G. and S.M.C. developed the purification processes and assessed quality by RPLC, size exclusion chromatography and CD. J.R.B., N.C.D. and K.R.L. developed and performed product-specific ELISAs. K.R.L., K.A.S., N.J.M. and J.R.B. engineered the strains used in production. J.J.C., N.A.C., D.L., C.A.M., C.B.M., N.J.M. and K.R.L. contributed to development of upstream processes. A.C. and L.B. developed the purification process for hGH. J.C.L., S.M.C., R.D.B. and W.S.H. designed the experimental strategy, supervised analysis and wrote the manuscript. All authors reviewed the manuscript.

1. Adiga, R. *et al*. Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-018-0259-1 (2018).
2. Boles, K.S. *et al*. Digital-to-biological converter for on-demand production of biologics. *Nat. Biotechnol.* **35**, 672–675 (2017).
3. Pardee, K. *et al*. Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248–259.e12 (2016).
4. Perez-Pinera, P. *et al*. Synthetic biology and microbioreactor platforms for programmable production of biologics at the point-of-care. *Nat. Commun.* **7**, 12211 (2016).
5. Dolsten, M. & Søgaard, M. Precision medicine: an approach to R&D for delivering superior medicines to patients. *Clin. Transl. Med.* **1**, 7 (2012).
6. Anonymous. Patient-centered drug manufacture. *Nat. Biotechnol.* **35**, 485 (2017).
7. Schellekens, H., Aldosari, M., Talsma, H. & Mastrobattista, E. Making individualized drugs a reality. *Nat. Biotechnol.* **35**, 507–513 (2017).
8. Love, K.R. *et al*. Comparative genomics and transcriptomics of *Pichia pastoris*. *BMC Genomics* **17**, 550 (2016).
9. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* **98**, 5301–5317 (2014).
10. Hamilton, S.R. Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**, 1441–1443 (2006).
11. Konstantinov, K.B. & Cooney, C.L. White paper on continuous bioprocessing. May 20–21 2014 Continuous Manufacturing Symposium. *J. Pharm. Sci.* **104**, 813–820 (2015).
12. Genentech Inc. Nutropin prescribing information. https://www.gene.com/download/pdf/nutropin_aq_prescribing.pdf (2016).
13. Center for Biologics Evaluation and Research & Center for Drug Evaluation and Research. Guidance for industry: for the submission of chemistry, manufacturing, and controls information for a therapeutic recombinant DNA-derived product or a monoclonal antibody product for in vivo use. https://www.fda.gov/downloads/biologicsbloodvaccines/guidancecomplianceregulatoryinformation/guidances/general/ucm173477.pdf (1996).
14. The European Agency for the Evaluation of Medicinal Products. CPMP position statement on DNA and host cell proteins (HCP) impurities, routine testing versus validation studies. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003322.pdf (1997).
15. Jawa, V. *et al*. Evaluating immunogenicity risk due to host cell protein impurities in antibody-based biotherapeutics. *AAPS J.* **18**, 1439–1452 (2016).
16. World Health Organization. Guidelines on the quality, safety, and efficacy of biotherapeutic protein products prepared by recombinant DNA technology. http://www.who.int/biologicals/biotherapeutics/rDNA_DB_final_19_Nov_2013.pdf (2013).
17. Canova-Davis, E. *et al*. Properties of a cleaved two-chain form of recombinant human growth hormone. *Int. J. Pept. Protein Res.* **35**, 17–24 (1990).
18. Merck & Co. Inc. Intron A product information 1–39 https://www.merck.com/product/usa/pi_circulars/i/intron_a/intron_a_pi.pdf (1986).
19. Matthews, C.B. *et al*. Reexamining opportunities for therapeutic protein production in eukaryotic microorganisms. *Biotechnol. Bioeng.* **114**, 2432–2444 (2017).
20. Timmick, S.M. *et al*. An impurity characterization based approach for the rapid development of integrated downstream purification processes. *Biotechnol. Bioeng.* **115**, 2048–2060 (2018).
21. Reinl, S.J. & Pogue, G.P. C-terminally truncated interferon. US Patent US20090025106A1 (2011).
22. Hermeling, S. *et al*. Structural characterization and immunogenicity in wild-type and immune tolerant mice of degraded recombinant human interferon alpha2b. *Pharm. Res.* **22**, 1997–2006 (2005).
23. Gibson, S.J. *et al*. N-terminal or signal peptide sequence engineering prevents truncation of human monoclonal antibody light chains. *Biotechnol. Bioeng.* **114**, 1970–1977 (2017).
24. Amgen Inc. Neupogen prescribing information. https://www.accessdata.fda.gov/drugsatfda_docs/label/2013/103353s5157lbl.pdf (2015).
25. Krishnan, S. *et al*. Aggregation of granulocyte colony stimulating factor under physiological conditions: characterization and thermodynamic inhibition. *Biochemistry* **41**, 6422–6431 (2002).
26. Lu, H.S. *et al*. Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony-stimulating factor: effect on stability and biological activity. *Arch. Biochem. Biophys.* **362**, 1–11 (1999).
27. Bönig, H. *et al*. Glycosylated vs non-glycosylated granulocyte colony-stimulating factor (G-CSF)—results of a prospective randomised monocentre study. *Bone Marrow Transplant.* **28**, 259–264 (2001).
28. Sörgel, F. *et al*. Comparability of biosimilar filgrastim with originator filgrastim: protein characterization, pharmacodynamics, and pharmacokinetics. *BioDrugs* **29**, 123–131 (2015).
29. Adamo, A. *et al*. On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* **352**, 61–67 (2016).
30. US Food and Drug Administration. Points to consider in the manufacture and testing of monoclonal antibody products for human use (US Government Publishing Office, 1997).

ONLINE METHODS

Generation of product-secreting strains. Wild-type *Komagataella phaffii* (NRRL Y-11430) was modified to express human growth hormone (rhGH), interferon alfa-2b (rIFNα-2b) or granulocyte-colony stimulating factor (rG-CSF) using sequences provided in **Supplementary Table 1**, codon-optimized for *P. pastoris* (GeneOptimizer service; Thermo Fisher Scientific, Waltham, MA) and cloned into a pPICZα-family vector (Thermo Fisher Scientific, Waltham, MA) flush with a truncated form of the *Saccharomyces cerevisiae* α-factor secretion signal under the immediate control of the methanol-inducible AOX1 promoter.

InSCyT system design and operation control. *Production module design and operation control.* Cultivation was performed in custom-modified Multifors 2 NW70 benchtop bioreactors equipped with 0.75 L flat-bottomed glass vessels (i.d. = 70 mm, height = 195 mm) (Infors USA, Annapolis Junction, MD). Magnetically coupled impellers enabled mixing (two six-blade Rushton or one Rushton and one three-blade marine).

Filtered (0.2 μm PTFE; Tisch Scientific, North Bend, OH) medical-grade oxygen (Airgas, Radnor, PA) was delivered through a metal sparge ring at the bottom of the vessel. Dissolved oxygen tension was measured using a 225 mm VisiFerm probe (Hamilton, Reno, NV) and controlled using a split-range proportional-integral (PI) controller, manipulating oxygen sparge rates (at low oxygen uptake rates) or stirrer speed (at high oxygen uptake rates).

A two-level cascade controller was used for temperature control. In the outer loop, vessel temperature was measured using the dissolved oxygen sensor's built in thermocouple and a PI controller was used to set the jacket set point. In the inner loop, the jacket temperature was measured using an embedded thermocouple and used to determine the duty cycle of the resistive heater or chilled 1:1 ethylene glycol:water loop (Julabo USA, Allentown, PA).

pH was measured with a model F-635 FermProbe (Broadley-James, Irvine, CA) and output-isolated pH transmitters (Hanna Instruments, Woonsocket, RI) and controlled using a deadband controller. Potassium hydroxide or phosphoric acid (5.0 M; Sigma-Aldrich, St. Louis, MO) were dosed into the bioreactor as needed by a four-channel, 12-roller Ismatec Reglo ICC peristaltic pump (Cole-Parmer, Vernon Hills, IL) through Ismatec 1.52 mm i.d. three-stop PharMed BPT tubing cassettes (Cole-Parmer, Vernon Hills, IL).

Disposable gamma-irradiated tubing assemblies were custom-built for the InSCyT system (High Purity New England, Smithfield, RI) to deliver media and collect perfusate. These assemblies were constructed of 1.52 mm i.d. three-stop PharMed BPT peristaltic cassettes, platinum-cured silicone tubing and HDPE carboys, along with appropriate PVDF barbed fittings. Flow into and out of the bioreactor was driven by Ismatec Reglo ICC peristaltic pumps with Asco Scientific three-way pinch valves (Cole-Parmer, Vernon Hills, IL). Clarified perfusate was withdrawn from the bioreactors using two custom modified probes each holding two porous 0.2-μm ceramic membranes (FISP, Flownamics, Madison, WI). Feed bottles were suspended on scales (Mettler Toledo, Columbus, OH) and closed-loop control was used to ensure constant flow. A conducting 3/16-inch grade 2 titanium probe (McMaster-Carr, Robbinsville, NJ), coupled to an Omron Automation 24V AC/DC monitoring relay (Allied Electronics, Fort Worth, TX) was used for level sensing and transmitting, activating the perfusion pump upon liquid contact. The level was maintained through equal media addition and perfusion rates. A periodic backwash was used to prevent membrane fouling.

Bioreactors were sampled automatically using a Seg-Flow 4800 Sampling System, a FlowFraction 400 and Seg-Mod modules (Flownamics, Inc., Madison, WI), and held at 4 °C until further analysis.

Perfusate adjustment module design and operation control. A pH adjustment module (pHAM) was used to adjust the pH of the perfusate before loading onto the first chromatography column. Supernatant was collected in a 1-L surge tank to balance flow rates between the bioreactor and the first column. Custom conductivity-based level sensors enabled automated startup of the downstream process at sufficient volume. Addition of adjustment solution in an in-line mixer (Stamixco, Wollerau, Switzerland) was used to adjust the perfusate pH before the first column. The pH was measured using a custom in-line pH probe (Van London Co., Houston, TX), and a PI controller determined the adjustment rate.

Purification module design and operation control. Up to three product-specific chromatography columns, operated in either bind and step gradient elution or flow-through mode, were used for purification. Flow was provided by a microannular gear pump (mzr-2905; HNP Mikrosysteme, Schwerin, Germany) and a flow sensor (SLI-2000; Sensiron, Zurich, Switzerland) in closed-loop PID control and passed through a debubbler/degasser (9000-1545, Idex Health and Science, Oak Harbor, WA). The columns were operated either independently or in series using multi-port (C65-3180IA; VICI Valco, Houston, TX) and solenoid (100T3MP24-62-5; BioChem Fluidics, Boonton, NJ) valves. Purification processes were operated using a predetermined sequence of steps, controlled either by time or by A_{280} measurements (model 280; Spectrum Labs, CA). Purified drug substance eluted from the final chromatography column was directed to the retentate reservoir within the formulation module.

Formulation module design and operation control. A tangential flow filtration system (TFF) (KRIIi; Spectrum Labs, CA) was used to concentrate and/or buffer exchange the eluted drug substance. The system was equipped with WaterSep Discover24 (or Discover12) membranes (5 kDa MWCO, 1.0 mm i.d.) (Marlborough, MA). Automated processing was enabled through custom scripting.

Module integration and automation. Modicon M221 PLCs (Schneider Electric, Andover, MA) were used for the connection of the thermocouple, motors, solenoids, level transmitters, UV transmitters and pH transmitters with the process local area network (LAN). The dissolved oxygen probes and oxygen mass flow controllers were connected to the process LAN via a Modbus to Ethernet endpoint (Sealevel, Liberty, SC), while the peristaltic pump and multi-port valve drives were connected to the process LAN via a RS-232 to Ethernet endpoint (Sealevel, Liberty, SC).

Wonderware (Lake Forest, CA) was used as a human–machine interface (HMI) to the integrated system. Custom scripts were written using Intouch QuickScript (Wonderware, Lake Forest, CA) to implement the USP, pHAM and DSP control loops, recipes, and operating sequences. These scripts were written in house or with assistance from Superior Controls (Seabrook, NH). Wonderware was also used as a data historian, with local download and processing performed using custom scripts in Matlab (MathWorks, Natick, MA).

Production of biologics using the InSCyT system. All numbered buffers referenced below are listed in **Supplementary Table 2**. All chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO).

Production of hGH using the InSCyT system: first-generation production. Bioreactors were filled with buffered glycerol-complex medium (BMGY) through pleated polyethersulfone 0.2-μm filters (Polycap 36 TC; GE Healthcare, Boston, MA). BMGY and buffered methanol-complex medium (BMMY) were sequentially fed to the reactor for outgrowth and induction, respectively[31]. Sigma A204 anti-foam (A8311; Sigma-Aldrich, St. Louis, MO) was added to BMMY at a concentration of 0.02% v/v. The working volume for all fermentations was 420 mL. Temperature and pH were maintained at 25 °C and 6.5, respectively.

BMGY was fed at 0.5 mL/min for 32 h to accumulate biomass. Dissolved oxygen was maintained at 100%. Mixing was achieved using two Rushton impellers. Perfusate was directed to the waste. After 32 h BMGY was automatically substituted for BMMY and perfusate was directed to the waste for 8 h. Forty hours after inoculation, perfusate was automatically diverted to the pHAM.

A first-generation pHAM was used to adjust the perfusate to pH 5.5 before loading onto the first chromatography column. pH was measured using an in-line pH probe (InLab Reach 425; Mettler Toledo, Columbus, OH) under a recirculation loop. A feedback controller was used to continuously vary the rate of phosphoric acid (500 mM) addition. A magnetic stir bar was used for mixing.

For each cycle of the purification module, 100 column volumes (CV) of adjusted perfusate were loaded onto a 5-mL prepacked multimodal cation exchange (MMCEX) column (CMM HyperCel; Pall Corporation, Port Washington, NY), equilibrated with buffer 1, washed with buffer 2, and eluted with buffer 3. Eluate from column 1 above 15 mAU was flowed through a 1-mL prepacked anion exchange (AEX) column (HyperCel STAR AX; Pall Corporation, Port Washington, NY). Flow-through from column 2 above

15 mAU was collected for formulation. The columns were stripped with buffer 4 and re-equilibrated with buffers 1 and 3, respectively.

Eluate from the final column was dialyzed against buffer 5 using 3.5K MWCO Slide-A-Lyzer G2 dialysis cassettes (ThermoFisher Scientific, Waltham, MA) according to the manufacturer's recommended protocol.

Production of hGH using the InSCyT system: second-generation extended production. The second-generation extended production of hGH was conducted the same as the first-generation production of hGH except for the following changes. BMGY was substituted for rich defined medium[31] containing 4% glycerol and BMMY was substituted for rich defined medium[31] containing 3% methanol. Dissolved oxygen was maintained at 25%. The pHAM described under "Perfusate adjustment module design and operation control" above was used. The adjustment fluid was 100 mM phosphoric acid. 80 CVs of adjusted perfusate (60 CVs for the purification cycle on day 10) was loaded onto the first column. Flow-through from the second column above 15 mAU was directed to the TFF module for formulation.

The formulation module was automatically triggered by the attached process computer at a fixed, regular interval, processing any eluate that had collected in the retentate reservoir. The feed pump was operated at a sufficient rate to maintain 50 mL/min crossflow velocity. Permeation rate was controlled though the use of a backpressure regulator valve set to maintain 30 psi transmembrane pressure. Concentration was performed on the TFF module (concentration factor 2.75) and then diafiltration was performed with buffer 6 (8 diavolumes).

Production of IFNα-2b using the InSCyT system: IFNα-2b process development experiment 1 (PDE 1). Production of IFNα-2b during PDE 1 was conducted the same as the first-generation production of hGH except for the following changes. Reactors were inoculated with an rIFNα-2b-secreting strain. 100 mM citric acid was used to adjust the pH to 5.0 in the pHAM. No recirculation loop was used in the pHAM; the pH was measured in tank by a model F-635 FermProbes (Broadley-James, Irvine, CA).

60 CV of supernatant was loaded onto a 5-mL prepacked MMCEX column (Capto MMC ImpRes; GE Healthcare Bio-Sciences, Pittsburgh, PA), equilibrated with buffer 7, washed with buffer 8 and eluted with buffer 9. Eluate from column 1 above 15 mAU was loaded onto a 5-mL hydrophobic charge induction chromatography (HCIC) column (MEP HyperCel; Pall Corporation, Port Washington, NY), equilibrated with buffer 9, washed with buffer 10 and eluted with buffer 11. Eluate from column 2 above 15 mAU was loaded onto a 5 mL prepacked cation exchange (CEX) column (SP Sepharose High Performance; GE Healthcare Bio-Sciences, Pittsburgh, PA), equilibrated with buffer 11, washed with buffer 12 and eluted with buffer 13. Eluate from column 3 above 10 mAU was collected for formulation. Columns were then stripped with buffer 4 and re-equilibrated at the equilibration conditions given above.

Eluate from the final column was dialyzed against buffer 14 using 3.5K MWCO Slide-A-Lyzer G2 dialysis cassettes (ThermoFisher Scientific, Waltham, MA) according to the manufacturer's recommended protocol.

IFNα-2b process development experiment 2 (PDE 2). Production of IFNα-2b during PDE 2 was conducted the same as production of IFNα-2b during PDE 1 except the elution buffer for the third column was buffer 15.

IFNα-2b process development experiment 3 (PDE 3). Production of IFNα-2b during PDE 3 was conducted the same as the second-generation production of hGH except for the following changes. Reactors were inoculated with an rIFNα-2b-secreting strain. The methanol-containing medium had 1% methanol. Dissolved oxygen was maintained at 40%. Mixing was achieved using one Rushton impeller (top) and one marine impeller (bottom). Perfusate pH was adjusted to 5.0 using 100 mM citric acid in the pHAM.

Purification was the same as in IFNα-2b PDE 2 except all columns were 1 mL and 110 CVs of supernatant were loaded onto the first column.

IFNα-2b process development experiment 4 (PDE 4). Production of IFNα-2b using the InSCyT system during PDE 4 was conducted the same as PDE 3 except for the following changes. Column 1 was eluted with buffer 16. Eluate from column 1 above 15 mAU was flowed through a 1 mL AEX column (HyperCel STAR AX; Pall Corporation, Port Washington, NY). Flow-through from column 2 above 15 mAU was loaded onto a 1 mL prepacked multimodal column (HEA HyperCel; Pall Corporation, Port Washington, NY), equilibrated with buffer 16, washed with buffer 17 and eluted with buffer 18. Eluate from column 3 above 10 mAU was collected for formulation and dialyzed as described in IFNα-2b PDE 1. Columns were then stripped with buffer 4 and re-equilibrated at the equilibration conditions given above.

IFNα-2b process qualification run. Production of IFNα-2b using the InSCyT system during the process qualification run was conducted the same as the second-generation production of hGH except for the following changes. Reactors were inoculated with an rIFNα-2b secreting strain. Upstream samples were not taken between hours 80 and 112 due to failure of the automated sampling system. For yield calculations, the titer during those hours is assumed to be the same as the titer of the previous sample pool (hours 56–78). Perfusate was adjusted to pH 5.0 in the pHAM using 100 mM citric acid.

Purification was the same as in IFNα-2b PDE 2 except 80 CVs of supernatant were loaded onto the first column. Eluate from column 3 above 10 mAU was directed to the TFF for formulation. No concentration was performed. Diafiltration was performed with buffer 19.

Production of G-CSF using the InSCyT system. Production of G-CSF was conducted the same as the second-generation production of hGH except for the following changes. Reactors were inoculated with an rG-CSF secreting strain. 0.1% CHAPs (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate) was added to the medium. Mixing was achieved using one Rushton impeller (top) and one marine impeller (bottom). Perfusate was adjusted to pH 5.0 in the pHAM using 100 mM citric acid .

100 CVs of adjusted perfusate were loaded onto a 5 mL prepacked MMCEX column (Capto MMC ImpRes; GE Healthcare Bio-Sciences, PA), equilibrated with buffer 7, washed with buffer 20 and eluted with buffer 21. Eluate from column 1 above 15 mAU was flowed through a 1 mL prepacked AEX column (HyperCel STAR AX; Pall Corporation, Port Washington, NY). Flow-through from column 2 above 12 mAU was loaded onto a 5 mL HCIC column (MEP HyperCel; Pall Corporation, Port Washington, NY), equilibrated with buffer 21, washed with buffer 22 and eluted with buffer 11. Eluate from column 3 above 12 mAU was directed to the TFF for formulation. Each of the columns was then stripped with buffer 4 and re-equilibrated with the equilibration conditions mentioned above for the first and third column and buffer 21 for the second column. No concentration was performed. Diafiltration was performed with buffer 23.

Analytical methods. *Wet cell weight.* To determine wet cell weight (g/L), 300 µL of resuspended bioreactor sample was dispensed into a preweighed Spin-X centrifuge tube (Corning, NY) and centrifuged at 15,000*g* for 10 min in technical triplicate. The mass of the pellet was determined by weight.

SDS–PAGE. SDS–PAGE was carried out under reducing conditions using Novex 12% Tris-Glycine Midi Gels (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's recommended protocol and stained using Instant Blue Protein Stain (Thermo Fisher Scientific, Waltham, MA).

Product-specific enzyme-linked immunosorbent assay (ELISA). Protein concentrations were determined using sandwich ELISA for G-CSF and direct ELISA for IFNα-2b as described elsewhere except for the following changes[32]. **Supplementary Table 3** shows the antibodies used. All reagents were added at 100 µL/well instead of 50 µL/well. Wash buffer was PBS, 0.05% Tween. Blocking buffer was PBS, 0.05% Tween, 0.25% BSA, and blocking steps were incubated at 37 °C for 1 h. Detection enzyme was incubated for 30 min. Substrate solution was ABTS (Thermo Fisher Scientific, Waltham, MA) and was incubated for 20 min. Stop solution was 0.1 M citric acid, 0.01% sodium azide. Plates were analyzed for absorbance at 410 nm/540 nm using a Tecan Infinite M200 Pro plate reader.

Process-related impurity analysis. Samples were analyzed for host-cell protein content using the *Pichia pastoris* 2nd Generation HCP ELISA kit from Cygnus Technologies (Southport, NC) according to the manufacturer's recommended protocol. Samples were analyzed for residual host-cell DNA using the resDNASEQ Quantitative *Pichia pastoris* DNA kit (Applied Biosystems, Foster City, CA) according to the manufacturer's recommended protocol. DNA was first extracted from each sample using the PrepSEQ Residual DNA Sample Preparation kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's recommended protocol. qPCR reactions were performed using a Roche LightCycler 480II equipped with LightCycler software release 1.5.0SP4 (Roche Molecular Systems, Inc., Indianapolis, IN). Analysis was performed using the built-in software as recommended in the manufacturer's protocol.

Isoelectric focusing (IEF). IEF was done using Novex pH 3-10 IEF gels (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's recommended protocol and stained with SimpleBlue SafeStain.

Chromatographic analyses. Reversed-phase liquid chromatography (RPLC) (for sample quantification and purity analysis) and size exclusion chromatography (SEC) (for quantification of high-molecular-weight species) were performed for hGH and G-CSF as described previously[20]. RPLC and SEC were carried out the same for IFNα-2b except RPLC operating conditions for IFNα-2b can be found in **Supplementary Table 4**, where buffers A and B are as described previously[20] and SEC running buffer for IFNα-2b was 50 mM ammonium bicarbonate, 200 mM arginine HCl, 0.02% sodium azide at pH 7.0. Column temperature was 60 °C. All chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO).

Liquid chromatography–mass spectrometry (LC–MS). 100 µg were used from each hGH and G-CSF sample for analysis. 25 µg were analyzed for IFNα-2b. hGH and IFNα-2b samples were dialyzed against 50 mM ammonium bicarbonate at pH 7.0 to a final concentration of about 1 mg/mL. G-CSF samples were dialyzed against 50 mM ammonium bicarbonate at pH 8.0 to a final concentration of about 1 mg/mL. Reference materials were treated the same as the samples. Samples were then transferred to an Amicon filter and spun at 5,000g for 15 min. hGH and IFNα-2b samples were digested with trypsin and G-CSF samples were digested with GluC/LysC using 1 µg (protein-to-enzyme ratio 50:1) of the respective enzyme and incubated overnight at 37 °C.

LC–MS equipment was used as described previously[33], except a microspray ion source was used. Mobile phases A and B were as described previously[33] and the flow rate was 200 µL/min. The gradient was as follows: 0–2 min 2% B with curve level 5, 2–30 min to 40% B, 30–39 min to 60% B, 39–42 min to 85% B until 47 min, 48–52 min to 2% B again. The gradient curve level was 6 from 2 to 52 min.

For peptide identification, raw data were searched against the product sequence using Thermo BioPharmaFinder 2.0 (Thermo Fisher Scientific). Peptide mass accuracy was set to 20 p.p.m. Oxidation of methionine residues and deamidation of asparagine residues were set as potential dynamic modifications. Final confirmation of the peptide identification was performed by manual inspection, extracting the base peak from the chromatogram and matching the MS–MS fragmentation data with the theoretical prediction.

Cell-based potency assays. Potency assays for G-CSF and hGH were conducted by Bioassay GmbH (Heidelberg, Germany). Cell-based proliferation assays for bioactivity determination of G-CSF samples were conducted according to Pharm. Eur. 01/2009:2206. Cell-based proliferation assays for bioactivity determination of hGH samples were conducted using NB2-11 cells and were compared to a standard (WHO NIBSC 98/574). Potency assays for IFNα-2b were conducted by Charles River Biopharmaceutical Services GmbH (Erkrath, Germany). *In vitro* cell-based assays for bioactivity determination of IFNα-2b samples were conducted according to Pharm. Eur. monograph 1110.

Circular dichroism (CD). A Jasco 815 spectrometer was used for CD. Spectra were recorded at a scanning speed of 200 nm/min, a bandwidth of 1 nm and an average of four scans. Near-UV CD spectra were recorded from 240 to 350 nm in a 10-mm path length cuvette, and far-UV CD spectra were recorded from 200 to 250 nm in a 1-mm path length cuvette. For hGH and IFNα-2b, samples were prepared by dialysis into a 10 mM sodium phosphate buffer at pH 6.75 with 0.1 g/L Tween 80 for IFNα-2b and pH 6.0 with 2.0 g/L Tween 20 for hGH using a 7K MWCO Slide-A-Lyzer G2 dialysis cassette. For G-CSF, samples were diluted with 10 mM sodium phosphate buffer at pH 4.4. Samples for near and far UV spectra were collected at approximate concentrations of 1.0 mg/ml and 0.1 mg/ml, respectively. Normalization was performed using concentration of the samples determined with a Hitachi U2910 UV-Vis spectrophotometer.

Nonclinical studies. *Material preparation.* InSCyT G-CSF batch 1 was used for all nonclinical studies. Endotoxin removal was performed using Pierce High-Capacity Endotoxin Removal Resin Spin Columns (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's protocol, except incubation was performed at 4 °C with gentle end-over-end mixing approximately once every 30 min for 3 h. Endotoxin removal was confirmed using the Pierce LAL Chromogenic Endotoxin Quantitation Kit (Thermo Fisher Scientific, Waltham, MA). Samples were filtered using a 0.2-µm filter (Thermo Fisher

Scientific, Waltham, MA). G-CSF concentration was determined using A_{280} measurements (DeNovix DS-11, Denovix, Wilmington, DE) after filtration. Samples were diluted in sterile, endotoxin-free 10 mM glutamic acid, 5% (w/v) sorbitol, at pH 4.4 to obtain a final concentration of 35 µg/mL (low dose) or 150 µg/mL (high dose) and sterilely aliquoted into single doses. Aliquots were stored at −80 °C and thawed immediately before administration.

Neupogen was purchased from Myoderm (Norristown, PA) and stored at 4 °C. Neupogen was diluted in sterile 10 mM sodium acetate, 5% (w/v) sorbitol, at pH 4.0 to 35 µg/mL immediately before administration. Sterile 10 mM sodium acetate, 5% (w/v) sorbitol, at pH 4.0 was used as a vehicle control.

Single-dose pharmacokinetics and pharmacodynamics study in Sprague Dawley rats. Pharmacokinetic profile and pharmacodynamics effect (neutrophil stimulation) of InSCyT G-CSF were evaluated by Toxikon Corporation (Bedford, MA) at two concentrations relative to a predicate control (Neupogen). Thirty-nine male Sprague Dawley rats made up four groups (randomized using Research Randomizer version 4.0 (Middletown, CT)), with 3 animals in untreated control group 1 and 12 animals in each of groups 2–4. Groups 2 and 3 received InSCyT G-CSF (115 µg/kg and 575 µg/kg, respectively) and group 4 received Neupogen (115 µg/kg). The InSCyT test articles and Neupogen control were administered once subcutaneously dorsally between the shoulders at the start of the study (day 1). Observations were conducted beginning before administration and throughout the study. Observations included, but were not limited to, changes in the skin, fur, eyes and mucous membranes, respiratory system, circulatory system, autonomic central nervous system, somatomotor activity, locomotor activity and behavioral pattern. Particular attention was paid to changes at the injection site. All animals survived for the duration of the study and were humanely euthanized via carbon dioxide inhalation at the end of the in-life portion (day 6).

Blood samples (approximately 0.5 mL) were collected for pharmacokinetic analysis before the dose and 0.5, 1, 2, 4, 8, 12, 24, 48, 72, 96 and 120 h after the dose from 3 animals per time point per test article group and predicate control group into tubes containing tripotassium EDTA. Tubes were placed on wet ice immediately following collection and centrifuged at 1,500g for 10 min. The processed plasma samples from all pharmacokinetic study animals were analyzed for test article and predicate control article concentration using an ELISA method. Briefly, the assay was developed based on a commercial kit (Quantikine ELISA) specific for recombinant human G-CSF in solution. The assay employed the quantitative sandwich enzyme immunoassay technique. A monoclonal antibody specific for G-CSF was precoated onto a microplate. G-CSF standards and samples were allowed to bind to the immobilized antibody. After washing, an enzyme-linked polyclonal antibody specific for G-CSF was added for detection of the bound G-CSF. Following a wash to remove any unbound antibody-enzyme reagent, enzyme substrate solution was added to the wells. The color produced was directly proportional to the concentration of G-CSF. The color intensities were measured using a microplate reader. The concentration of the test articles was determined from a standard curve. Data analysis of the resulting plasma concentrations of InSCyT G-CSF test article and Neupogen predicate control was performed using WinNonlin software v6.3. A noncompartmental analysis was performed using NCA model 202 for pharmacokinetic parameter determination. AUC values were calculated using the trapezoidal linear interpolation method. Concentration values below the lower limit of quantification of 39 or 78 pg/mL, as applicable, were set to zero for analysis. $t_{1/2}$ was calculated from the slope of the curve for 4,8 and 12 h for InSCyT G-CSF and 8, 12 and 24 h for Neupogen.

Blood samples (approximately 2.0 mL) were separately collected 24 h after the dose into tubes containing dipotassium EDTA for neutrophil analysis from three rats from each group. Whole blood was analyzed for neutrophil count.

Five-day repeat dose study in Sprague Dawley rats. Toxicology of InSCyT G-CSF was evaluated as compared to a predicate control (Neupogen) in Sprague Dawley rats when administered at 115 µg/kg subcutaneously dorsally once per day for 5 d. Thirty Sprague Dawley rats made up three groups (randomized using Research Randomizer version 4.0 (Middletown, CT)), with five animals per sex in each group. Group 1 received a vehicle control, group 2 received InSCyT G-CSF and group 3 received Neupogen. Toxicity was evaluated on the basis of survival, clinical signs, body weight, quantitative food consumption, hematology, serum chemistry, organ weights and macroscopic findings.

Body weights and food consumption were measured daily. Observations were performed twice daily and included, but were not limited to, changes in the skin, fur, eyes and mucous membranes, respiratory system, circulatory system, autonomic central nervous system, somatomotor activity, locomotor activity and behavioral pattern. Particular attention was paid to changes at the injection site. All animals survived the study without observed toxicity of any kind and were humanely euthanized via carbon dioxide inhalation at the end of the in-life portion (day 6) of the study.

Clinical pathology analysis was performed on blood samples obtained before necropsy on day 6 from all animals (approximately 24 h after the last dose). Hematology parameters assessed included red blood cell count, hemoglobin, mean corpuscular volume, mean corpuscular hemoglobin concentration, differential white blood cell count, white blood cell count, hematocrit, mean corpuscular hemoglobin, platelet count and reticulocytes. Clinical chemistry parameters assessed included alanine aminotransferase, albumin, albumin/globulin ratio, alkaline phosphatase, aspartate aminotransferase, blood urea nitrogen, calcium, chloride, cholesterol, gamma glutamyltransferase, creatinine, globulin, glucose, phosphorus, potassium, sodium, total bilirubin, total protein and triglycerides.

Quantitative, continuous data from the study were analyzed using one-way ANOVA using Provantis 9.3.1. Differences between dose groups was considered statistically significant only if the probability of the differences being due to chance is equal to or less than 5% ($P < 0.05$).

Statistics. Sample sizes, error bars and types and number of replicates are defined in the figure legends or in the corresponding method sections above. The Kolmogorov–Smirnov test was used to determine significance for pharmacokinetics studies ($P = 0.9963$). For the toxicology studies, significance was determined using one-way ANOVA as described above. Differences were considered statistically significant only if $P < 0.05$.

Animal welfare. To the best of our knowledge, the nonclinical studies described here did not unnecessarily duplicate previous testing and there were no non-animal alternatives acceptable for the evaluation of the test article as defined by the protocol. No evidence of pain and distress was reported to the veterinarian or study director. Protocols for each study were approved by Toxikon's institutional animal care and use committee (IACUC). Toxikon strictly adhered to common standards in maintaining the animal care and use program[34–38].

Life Sciences Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. Custom code used in this study is available from the corresponding author upon reasonable request.

Data availability. The datasets generated and analyzed in this study are available from the corresponding author upon reasonable request.

31. Matthews, C.B., Kuo, A., Love, K.R. & Love, J.C. Development of a general defined medium for *Pichia pastoris*. *Biotechnol. Bioeng.* **115**, 103–113 (2018).
32. Hornbeck, P., Winston, S.E. & Fuller, S.A. Enzyme-Linked Immunosorbent Assays (ELISA). in *Curr. Protoc. Mol. Biol.* 11.2.1–11.2.22 (Wiley, 1991).
33. Wang, Y.A. *et al.* Integrated bottom-up and top-down liquid chromatography-mass spectrometry (LC-MS) for characterization of recombinant human growth hormone degradation products. *Anal. Chem.* https://doi.org/10.1021/acs.analchem.7b03026 (2017).
34. United States Department of Agriculture (USDA). *Animal and Plant Health Inspection Service. 9 CFR Ch. 1, Subchapter A - Animal Welfare* (US Government Publishing Office, 2013).
35. Office for Laboratory Animal Welfare (OLAW). Public Health Service policy on humane care and use of laboratory animals. *Health Research Extension Act of 1985* (US Government Publishing Office, 1985).
36. National Research Council. *Guide for the Care and Use of Laboratory Animals.* (National Academies Press, 2011).
37. *Biological Evaluation of Medical Devices - Part 2: Animal Welfare Requirements ISO 10993-2* (International Organization for Standardization, 2006).
38. Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC) International. https://www.aaalac.org/ (2018).

nature research

Corresponding author(s):   J. Christopher Love

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▶ Experimental design

1. Sample size

Describe how sample size was determined.

The experiments presented in this work were representative of greater than 35 independent operational replicates performed under similar conditions to determine optimal process parameters for production and their impact on quality attributes of proteins. Sample sizes were not predetermined. Instead, replicate experiments (min n=3) were initially performed for each condition or process. Variations or quality deviations were noted. Process conditions were adjusted accordingly and experiments were performed again in replicate to confirm new conditions. Iterations were performed until variation and quality was acceptable. Sample size for non-clinical studies was selected to be the minimum number of animals required to obtain statistically significant results.

2. Data exclusions

Describe any data exclusions.

Technical outliers were excluded from the cell-based potency assays for biologic activity of hGH (as determined by a third-party CRO Bioassay GmbH based on their Quality Assurance protocols for data generated in these assays and described in Methods). No data were excluded from analysis in other experiments reported.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The production and characterization of G-CSF, IFNalpha-2b, and hGH presented in this work were representative of more than 35 independent operational replicates performed with similar parameters to map optimal production conditions. Quality attributes of materials were consistent with expectations (for example, higher dissolved oxygen led to higher levels of oxidation of protein). Materials produced in this work were regularly generated in triplicate using three independent automated production systems.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

For the non-clinical studies presented in this work randomization was performed to allocate animals into experimental groups using Research Randomizer software, version 4.0.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

For the non-clinical studies presented in this work blinding was not performed. The study employed methodology to minimize uncertainty and to control bias for data collection and analysis, however, which included but was not limited to: concurrent control data, system suitability assessment, randomization, and method controls such as blanks and replicates.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

▶ Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Wonderware 2014 R2 was used as a human-machine interface to the integrated system and as a data historian. GraphPad Prism v7 was used to plot and analyze all data in the manuscript and supplemental material, except for the following. MATLAB 2017a was used for plotting and analysis of the fermentograms and UV traces presented. LightCycler software release 1.5.0SP4 was used for qPCR analysis. Empower 3 was used for control in chromatographic analyses. Thermo BioPharmaFinder 2.0 was used for analysis of LCMS data. Provantis 9.3.1 and WinNonlin 6.3 was used for analysis in the non clinical studies.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

▶ Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Custom materials used for operation of the InSCyT system are available from the authors or from the companies noted in the methods section.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

The antibodies used for the G-CSF ELISA include Biolegend BVD13-3A5 (Cat# 502102, Lot# B129744), Biolegend BVD11-37G10 (Cat# 506702, Lot# B191907), and Abcam Streptavidin-HRP ab7403 (Lot# GR305788-2). The antibodies used for the IFNalpha-2b ELISA include AssayPro 31168-05121 (Lot# IB071910405) and Abcam Streptavidin-HRP ab7403 (Lot# GR305788-2). The antibodies used for the G-CSF ELISA in the PK studies were from the commercial Quantikine kit (R&D Systems Cat# DCS50). For Biolegend products, each lot of antibody was quality validated by ELISA using recombinant G-CSF. For Abcam products, each lot was validated against Biotinylated IgG in a standard capture ELISA using a peroxidase substrate. For AssayPro products, each lot of antibody was validated by ELISA using biotinylated recombinant IFNaplha-2b. The Quantikine kit was validated by ELISA using recombinant human G-CSF.

10. Eukaryotic cell lines

   a. State the source of each eukaryotic cell line used.

> Komagataella phaffii NRRL Y-11430 was obtained from ATCC (Catalog number 76273).

   b. Describe the method of cell line authentication used.

> K. phaffii NRRL Y-11430 and derivatives described in the methods were authenticated by genome sequencing as reported in Love, K. R. et al. BMC Genomics (2016) 17:550.

   c. Report whether the cell lines were tested for mycoplasma contamination.

> Only yeast were used in this study and all cultivations used animal-free components; mycoplasma contamination/testing is not applicable and was not performed.

   d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

> No commonly misidentified cell lines were used.

▶ Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

> Thirty-nine Sprague Dawley rats (male) were used in PK/PD studies. All rats were at least 5 weeks old and at least 200g. Thirty Sprague Dawley rats (15 male and 15 female) were used in the repeated-dose study. Weights ranged from 232.7 - 328.0g. All females were non-pregnant and nulliparous.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

> No human research participants were involved in this study.