Modular development enables rapid design of media for alternative hosts

Andrew M. Biedermann[1,2], Isabella R. Gengaro[1,2], Sergio A. Rodriguez-Aponte[1,2], Kerry R. Love[1,2], J. Christopher Love[1,2]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[2]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 01239, United States

*Correspondence to: clove@mit.edu

Target journal: Biotechnology & Bioengineering
Article (<15 double-spaced pages of text and no more than 10 figures/tables)

Keywords: Media development; media blending; *Pichia pastoris*; alternative hosts; automation

Abstract

Developing media to sustain cell growth and production is an essential and ongoing activity in bioprocess development. Modifications to media can often address host or product-specific challenges, such as low productivity or poor product quality. For other applications, systematic design of new media can facilitate the adoption of new industrially relevant alternative hosts. Despite manifold existing methods, common approaches for optimization often remain time and labor intensive. We present here a novel approach to conventional media blending that leverages stable, simple, concentrated stock solutions to enable rapid improvement of measurable phenotypes of interest. We applied this modular methodology to generate high-performing media for two phenotypes of interest: biomass accumulation and heterologous protein production, using high-throughput, milliliter-scale batch fermentations of *Pichia pastoris* as a model system. In addition to these examples, we also created a flexible open-source package for modular blending automation on a low-cost liquid handling system to facilitate wide use of this method. Our modular blending method enables rapid, flexible media development, requiring minimal labor investment and prior knowledge of the host organism, and should enable developing improved media for other hosts and phenotypes of interest.

Introduction

Achieving high volumetric productivities of biologic drugs in cultivation is a key step in advancing candidate biologic drugs. The outcome of this effort ultimately impacts manufacturing costs as well as readiness for transitioning clinical-stage development (Love, Love, & Barone, 2012). The development of standard, chemically defined media for established manufacturing hosts, such as CHO, has made such transitions efficient for monoclonal antibodies by achieving high biomass accumulation, cell viability, operational consistency, and specific productivities, streamlining development efforts (McGillicuddy, Floris, Albrecht, & Bones, 2018; Rodrigues, Costa, Henriques, Azeredo, & Oliveira, 2012). Nonetheless, optimizing productivity or quality attributes for a specific product often still requires further refinement of media (Ritacco, Wu, & Khetan, 2018). Such development may require evaluating dozens of variants derived from a common standard formulation to address the specific challenges encountered (Gagnon et al., 2011; Loebrich et al., 2019). Media development for entirely new biomanufacturing technologies, such as alternative hosts (Matthews, Kuo, Love, & Love, 2017a) or new product modalities (Lu et al., 2016), may also require new formulations or extensive optimizations due to limited prior knowledge.

Common approaches to develop a medium to optimize a phenotype of interest are often labor intensive, low throughput, or rely heavily on extensive analytical capacity (Galbraith, Bhatia, Liu, & Yoon, 2018). For example, analysis of residual media after cultivation requires extensive capabilities for analytical characterization and prior experience with the manufacturing host to identify potentially limiting or toxic media components (Mohmad-Saberi et al., 2013; Pereira, Kildegaard, & Andersen, 2018). As a result, optimizations can be slow and iterative. Furthermore, for an alternative host such as *Komagataella phaffii* (formerly known as *Pichia pastoris*), there is substantially less, if any, prior knowledge available to establish profiles for residual components in media after fermentation. Other analytical techniques like RNA-seq combined with methods for reporter metabolite analysis can guide media optimization, to

75 generate testable hypotheses regarding beneficial modifications to media (Matthews et al.,
76 2017a). Such genome-scale approaches, however, require prior host-specific knowledge, such
77 as well-annotated genomes, and are still limited by slow iteration and labor-intensive
78 preparations of new media to test the hypotheses generated from computational analyses.
79 Alternative strategies for blending basal components for media allow linear combinations
80 of existing media to explore many variations rapidly (Jordan, Voisard, Berthoud, & Tercier,
81 2013). While this approach avoids slow iterative analyses, the typical experiment is labor
82 intensive to perform, often requiring independent preparations of over a dozen stock media to
83 combine (Rouiller et al., 2013). Similar to analytical-based approaches for optimization, the
84 selected variations of media are simultaneously guided and constrained by prior experience and
85 media designs, which may limit the breadth of components examined (Kennedy & Krouse,
86 1999). For less established hosts with fewer available formulations of media, media blending
87 may also require fully *de novo* formulations for initial studies. Further complicating such designs,
88 different and new components for media can present challenges in solubility or unanticipated
89 interactions with other elements in the formulations (Ritacco et al., 2018). New approaches to
90 blending could, however, enable fast, flexible experimentation and minimize the time, labor, and
91 analytical development needed initially to optimize media for new applications and phenotypes.
92 Here, we present a novel and generalizable approach for the modular development of
93 media and demonstrate its use to create optimized media for two different phenotypes—cellular
94 growth and recombinant expression of a protein (as measured by the secreted heterologous
95 protein titer) from *Pichia pastoris*. Our approach comprises two modular parts for blending and
96 optimization. We determined that a set of simple concentrated stock solutions constructed in
97 defined modules could generate many media by blending or dilution. We then automated a
98 simple, inexpensive liquid handling system (Opentrons OT-2) to enable high-throughput
99 screening for the effects of diverse media on a phenotype of interest in milliliter-scale batch
100 cultures. To maximize the benefit of this automated blending, we also developed an algorithmic

101 framework for systematic modular media optimization, beginning from a simple minimal media
102 (here a YNB-based one). This framework provides insights pertaining to key media components
103 during stages of optimization, as well as overall mapping of the design space for the media. In
104 the examples presented here, the resulting defined media developed with this strategy
105 outperformed commonly used BMGY and BMMY complex formulations for biomass
106 accumulation and secreted heterologous protein production.
107

108 **Materials and Methods**

109 *Strains and cultivation conditions*

110 Media for evaluating biomass accumulation were developed using a previously
111 described strain expressing G-CSF under control of the pAOX1 promoter (Crowell et al., 2020).
112 24-well plate screens were conducted as described previously, except cells were grown on a
113 labForce shaker and were only sampled 24 hours after inoculation (Matthews et al., 2017a).
114 BMGY, BMMY, and RDM media were formulated as described previously for shake flasks
115 (Matthews et al., 2017a). All cultivations were inoculated from a working cell bank at an initial
116 cell density of 0.1 OD/mL. For each working cell bank, cells grown in 1 L shake flasks with a 200
117 mL working volume of RDM were harvested during exponential growth (4-5.5 OD/mL) via
118 centrifugation at 1500 rcf for 4 minutes at 23 ^{o}C and resuspended in an equal volume mixture of
119 RDM and 50 v/v% glycerol. This mixture was then distributed into 700 μL aliquots and stored at
120 -80 ^{o}C, resulting in a cell density of ~30 OD/mL for the cell bank.

121 Media for evaluating enhanced production were developed using a strain expressing a
122 rotavirus-derived subunit vaccine candidate, P[8], under the control of the pAOX1 promoter
123 described previously (Dalvie et al., 2020). Biomass accumulation proceeded for 24 hours; cells
124 reached an initial induction density of ~20 OD/mL. Cultures were then exchanged into
125 production media and allowed to produce protein for an additional 24 hours. Supernatant was
126 harvested by centrifugation at 1500 rcf for 4 minutes at 23 ^{o}C and filtered using a Captiva 96

127 well 0.2 μm filter plate (Agilent Technologies, Santa Clara, CA) prior to titer measurement by
128 RP-UHPLC.
129 Media components and supplements were purchased from Sigma-Aldrich, St. Louis,
130 MO, unless otherwise indicated in the supporting information. A table of supplement and stock
131 solutions with screening concentrations is also included in the supporting information. During
132 modular optimization, all media were prepared in high throughput using an Opentrons OT-2
133 liquid handler (Opentrons, Brooklyn, NY, software version ≥3.16.1) using Openblend. Modular
134 media blending code and instructions for setup and operation are provided in the Openblend
135 package (https://github.mit.edu/lovelab/openblend). For consistency, media used in final head-
136 to-head comparisons were prepared in bulk and filter sterilized through a 0.2 μm benchtop filter.
137
138 *Analytical procedures*
139 Biomass was measured by optical density at 600 nm as described previously (Matthews
140 et al., 2017a). An Agilent Bravo liquid handler was used to dilute samples prior to
141 measurements of OD into the Tecan Infinite M200 Pro plate reader.
142 Reverse phase ultra-high performance liquid chromatography (UHPLC) analysis was
143 performed on Agilent 1290 Infinity II UHPLC system controlled using OpenLab CDS software
144 (Agilent Technologies, Santa Clara, CA). The concentration of protein was determined using a
145 Poroshell 120 SB-Aq column (2.1 x 50 mm, 1.9μm) operated at 1.0 mL/min and 70 ^{o}C (Agilent
146 Technologies, Santa Clara, CA). Buffer A was 0.1% (v/v) TFA in water and buffer B was 0.1%
147 (v/v) TFA, 0.5% (v/v) water in ACN. A gradient was performed as follows: 30% B for 1 min., 30-
148 40% B over 3 min., 40-90% B over 0.5 min., 90% B for 0.5 min., 90-30%B over 0.5 min., and
149 30% for 1 min.; total method run time was 6.5 minutes. Sample injection volumes were 50μL. A
150 diode array detector was set for absorbance detection at 214nm. Data analysis was completed
151 using OpenLab CDS Data Analysis (Agilent Technologies, Santa Clara, CA).

Statistical analysis and DOE design was conducted using JMP (SAS Institute, Cary NC). Quadratic models were fitted using effect screening and non-significant terms (adjusted p-value > 0.01) were eliminated sequentially in order of decreasing adjusted p-value to avoid overfitting. Data was plotted using Prism 8.4.0 (GraphPad Software, San Diego, CA).

Results

Design of approach for modular media blending

We sought to develop an approach capable of identifying important, beneficial modifications for media tailored to a given phenotype of interest. We reasoned that key requirements for such an approach would be that it is fast and automatable, with minimal dependence on complex analytical assay development. Such features would enable routine application to any measurable phenotype of interest. In general, media blending allows both speed and low analytical complexity. We aimed to retain these features while minimizing the labor and constraints on compositions imposed by linear combinations of fully formed and unique media. We reasoned that diverse and flexible blends of media could be created by defining simple concentrated stock solutions as basic modules to combine further. These modules would comprise individual components or common subsets of components with compatible solubilities (e.g. YNB). If media components could be formulated in concentrated stock solutions that could be stored stably over time, then the components could be routinely and interchangeably combined and diluted to the desired final concentrations. This approach would yield a broadly applicable modular strategy for media blending amenable to conventional liquid handling automation.

To test the feasibility of this approach, we first assessed whether many common media components could be formulated in concentrated stable aqueous stock solutions. Using the

178 CHO medium eRDF as a reference, we estimated the solubility of each component of this
179 medium, using data from AqSolDB as well as other online sources (Combs, 2012; FSA Panel
180 on Additives and Products or Substances used in Animal Feed (FEEDAP), 2011; Ritacco et al.,
181 2018; Schnellbaecher, Binder, Bellmaine, & Zimmer, 2019; Sorkun, Khetan, & Er, 2019;
182 Yamamoto & Ishihara, n.d.). We compared the estimated solubility of each media component to
183 its concentration in eRDF and found that, individually, most media components are soluble at
184 levels >10x higher than their eRDF concentration (**Figure 1A**). The existence of a wide range of
185 commercially available concentrated supplements further supports this result: >50x
186 concentrated solutions of amino acids, vitamins, lipids, and trace metal supplements are
187 common and commercially available.

188 Next, we used the product information of commercially available supplements, literature
189 sources, and inspection to estimate the percentage of eRDF media components that could be
190 stored in stable solutions for >6 months. We estimated that over 75% of eRDF components met
191 this criterion (**Figure 1B**). To address stability challenges caused by less stable components,
192 we reasoned that less stable components or supplements, such as vitamins, could be prepared,
193 aliquoted, and stored frozen for long-term storage (Schnellbaecher et al., 2019); these aliquots
194 could then be thawed and used within a defined period to mitigate component instability and
195 enable their integration into our modular blending strategy. Together, these solubility and
196 stability data suggested that a modular approach to media development could be defined in this
197 way to accommodate a range of new formulations easily.

198 We next automated the process for constructing media, using the Opentrons OT-2. We
199 chose this liquid handler due to its low cost, reliability, and compatibility with simple formats for
200 data input, such as Excel spreadsheets. We then created an open-source Python package,
201 named Openblend, which simplified the media construction process by handling routine
202 experimental design and execution steps (Figure 1C). Openblend creates an experimental
203 design spreadsheet, specifying the number of 24 well plates, the desired media composition of

204 each well, and stock solution names and concentrations. The script then checks the feasibility
205 of the experimental design, ensuring that the total volume of each well will not exceed the target
206 volume and avoiding the addition of sub-microliter stock solution volumes. If the design passes
207 this assessment, the script then outputs a new spreadsheet containing the setup for the OT-2
208 deck and required volumes of stock solutions, providing a user with instructions on how to setup
209 the OT-2 liquid handler. We found that our typical time to execute this script, setup the OT-2
210 and initiate plate building was ~15 minutes, and the time for the automated steps was about two
211 hours.

212 Finally, we defined a modular approach for optimization to effectively leverage the
213 Openblend tool (**Figure 1D**). Beginning from an initial basal medium, improved media are
214 constructed through successive rounds of optimization. In each round, a library of media
215 components and supplements are screened to identify beneficial additives. These additives are
216 then screened in combination and over a range of concentrations to further optimize the
217 performance of the medium. Each modular addition and optimization of additives can be guided
218 simply by measurements of the phenotype of interest (e.g. biomass accumulation). This greedy
219 approach to multi-dimensional optimization could continue iteratively until the resulting media
220 met desired specifications, all available media components were explored, or no additional
221 gains in performance realized.

222

223 *Application to Developing a Medium for Biomass Accumulation*

224 To assess the utility of this blending-based approach, we next aimed to identify and
225 optimize the concentration of media components beneficial for rapid biomass accumulation of *P.*
226 *pastoris* in batch cultivation. We previously described a rich defined medium (RDM) (Matthews
227 et al., 2017a), capable of high growth rates during biomass accumulation. One challenge
228 encountered with this formulation, however, was that precipitates can form at higher pH values
229 that require filtering during bulk preparations. Nonetheless, this medium provided a relevant

230 comparison for assessing the medium realized with our new approach due to its prior
231 demonstrated benefits relative to complex media. Following our modular approach, we
232 improved biomass accumulation by optimizing the accumulated optical density at 600 nm after
233 24 hours of cultivation.

234 Algorithms for optimizing systems based on multiple dimensions are often sensitive to
235 initial conditions used (Zakharova & Minashina, 2015). Given this potential confounding effect
236 here, we tested first the effects of the types of carbon source, nitrogen source, and pH set point
237 on biomass accumulation, using 1x YNB without amino acids or ammonium sulfate (YNB) to
238 satisfy minimum requirements for the concentrations of trace elements. We conducted a full-
239 factorial DOE using glycerol, glucose, and fructose as carbon sources; urea and ammonium
240 sulfate as nitrogen sources; and potassium phosphate as a buffer with pH values of 5, 5.75, and
241 6.5. We selected initial concentrations of 40 g/L, 4 g/L urea or the N-mol equivalent for
242 ammonium sulfate, and 10 g/L potassium phosphate, similar to values used in other media for
243 *Pichia pastoris* (Matthews et al., 2017a). A least squares regression model, including individual,
244 combination, and quadratic effects was fit to the log of optical density after 24 hours, a proxy
245 variable for the average growth rate (R^2 = 0.81). We determined that the two most significant
246 model terms were the type of carbon source and the interaction of the nitrogen source with pH
247 (**Figure 2A**). We found that cells grew significantly faster on metabolically related sugars
248 (glucose and fructose) than on the polyol (glycerol) commonly used for Pichia during biomass
249 accumulation (**Figure 2B**). This result affirms prior reports where glucose has been used for
250 biomass accumulation of Pichia (Guo et al., 2012; Moser et al., 2017).

251 The model also suggested that poor biomass accumulation during cultivation resulted
252 from a combination of ammonium sulfate as a source of nitrogen with low buffer pH (**Figure**
253 **2B**). This outcome may result from the production of acidic species associated with cellular
254 ammonium metabolism in the batch cultivation (Villadsen, 2015). Interestingly, the model
255 indicated slightly greater biomass was achieved with urea instead of ammonium sulfate. The

256 biomass accumulation of cultures grown with urea as a source of nitrogen were less sensitive to
257 reduced pH values (~5). We observed, however, that cultivations at pH 5 showed extensive
258 flocculation compared to those at 6.5. Given the insensitivity of urea-fed cultivations to buffer
259 pH and the high solubility and potential for low-cost sourcing of fructose, we therefore chose to
260 include fructose, urea, and a potassium phosphate buffer with a pH of 6.5 in our initial media
261 formulation.

262 With this basal formulation determined, we next screened for concentration-dependent
263 interactions of other key additives to the media and then optimized concentration-dependent
264 parameters. Following the same approach for screening effects, we conducted a full factorial
265 DOE over a broad range of media component concentrations: YNB (0.5, 1, 2x), fructose (10, 30,
266 50 g/L), urea (1, 4, 7 g/L), and potassium phosphate adjusted to a pH of 6.5 (4, 10, 16 g/L). The
267 resulting model identified fructose as a concentration-sensitive parameter (R^2=0.73) (**Figure**
268 **2D**). Terms involving the concentration of YNB were also highly ranked, but not statistically
269 significant. No significant interactions between components were identified in the model. We
270 therefore sought to better understand the concentration dependence of fructose and YNB
271 independently (**Figure 2E**), over an 8-fold range of concentrations. As expected, biomass
272 accumulation was highly sensitive to fructose concentration, with an optimum around 22.5 g/L of
273 fructose. The concentration of YNB had minimal effect on biomass accumulation; the presence
274 of trace elements supplied by YNB, however, was essential to growth. Based on these results,
275 we chose concentrations of 22.5 g/L fructose, 1x YNB, 7 g/L urea, and 10 g/L potassium
276 phosphate buffer. We reasoned that although biomass accumulation was relatively insensitive
277 to the concentrations of YNB and urea, higher concentrations could provide improved media
278 depth in future applications. We named this basal formulation DM1_dev0.

279 We next assessed what additional media components could improve biomass
280 accumulation. To test over 60 different components individually would require over 60 individual
281 solutions. Such an approach would scale linearly with new components; instead, we chose to

282 screen groups of related components, using commercially available pre-mixed supplements. We
283 compiled a library of 16 commercial supplements and industrially-relevant surfactants containing
284 more than 60 unique components and screened their individual effect on biomass accumulation
285 after 24 hours. In this way, we reasoned we could efficiently identify critical classes of
286 components related to the phenotype of interest and potentially deconvolve specific individual
287 additives of interest by inference. We used the recommended concentrations of each
288 supplement as supplied in product information, or critical micelle concentrations, and prior
289 knowledge for broad classes in yeast media to set reasonable screening concentrations
290 (Supporting Information). We identified five beneficial and two detrimental supplements that
291 significantly impacted biomass accumulation ($p_{adj} < 0.02$; 1-way-ANOVA) (**Figure 2F**). In
292 general, the results suggest that supplementation with amino acids and trace metals were
293 beneficial for accumulating biomass, while two surfactants, Tween 20 and CHAPS, were
294 detrimental. For this phenotype, the effects of vitamin and lipid supplements were minor;
295 supplements from either supplement category were not significantly beneficial or detrimental to
296 biomass accumulation. Our earlier experiments suggest that vitamins are essential but
297 concentration agnostic (**Figure 1E**), while lipid supplementation provides no clear benefit for
298 biomass accumulation.

299 Based on these results, we chose to test whether combinations of supplements of amino
300 acids and trace salts could yield synergistic improvements in biomass accumulation. We
301 screened pairwise combinations of the five beneficial supplements of mixed composition and
302 ranked the performance of our supplementation strategies (**Figure 2G**). A combination of 1x
303 MEM amino acids with 0.1 v/v% PTM1 salts resulted in the highest yield of biomass, though we
304 observed strong performance from other combinations of amino acid and trace metal
305 supplements. Based on these data, we chose to add MEM amino acids and PTM1 salts in our
306 basal medium and optimized their concentrations (**Figure 2H**).

307 Based on these results, we elected 0.1 v/v% PTM1 salts and 1x MEM amino acids, in
308 order to balance the moderate benefits and potentially high costs of amino acids. We found,
309 however, that the inclusion of the PTM1 salts in liter-scale preparations produced fine
310 precipitates, which can impede sterile transfers in use. To overcome this challenge, we
311 screened a broad range of PTM1 salts concentrations to identify the minimum concentration
312 required for improved outgrowth performance (**Figure 2I**). We found that PTM1 addition at
313 concentrations as low as 0.0005 v/v% led to increased biomass accumulation. We therefore
314 revised our PTM1 salts concentration to 0.01 v/v%, a concentration high enough to obtain the
315 benefits of PTM1 supplementation without inducing precipitate formation. This formulation we
316 named DM1.
317 Completing this series of optimizations with our iterative modular approach to define a
318 new formulation of medium, we then compared with other common media used to grow *P.*
319 *pastoris*. We evaluated the performance of this new optimized medium (DM1) relative to the
320 unsupplemented basal medium (DM1_dev0), the rich defined medium (RDM) we had previously
321 developed, and a common medium 4 v/v% glycerol BMGY. We found that DM1 yielded the
322 highest biomass accumulation, with significantly higher biomass accumulation relative to RDM
323 and BMGY (**Figure 2J**). This result demonstrates the utility of our modular strategy here for
324 media development that yielded an improved formulation for biomass accumulation compared
325 to other common media with minimal time and labor investment, and without requiring complex
326 analytical methods like mass spectrometry or RNA-sequencing.
327
328 *Identifying media conditions important to heterologous protein production in* K. phaffii
329 In addition to the time and labor savings of modular media development, our proof-of-
330 concept experiments demonstrated that this approach creates a flexible medium that can be
331 rapidly adapted to new growth phenotypes, as well as a data package that the identifies media
332 conditions important to the phenotype of interest. We reasoned that these additional benefits

333 could be particularly relevant for optimizing production of heterologous proteins. Understanding
334 which media components contribute most significantly to productivity could improve culture
335 performance and help identify important metabolic pathways or physiological effects for further
336 study.

337 To develop a medium for improved production of a recombinant protein, we chose to use
338 a strain engineered to secrete a rotavirus-derived subunit vaccine component, VP4-P[8], as a
339 model protein. We have previously demonstrated that this viral antigen can be expressed at
340 high titer under the control of the methanol-inducible pAOX1 promoter in BMMY media (Dalvie
341 et al., 2020). Similar to our initial approach to optimize a medium for growing biomass, we first
342 determined and optimized the concentrations of the sources for carbon and nitrogen, along with
343 the pH. The expression of P[8] in the strain tested uses the methanol-dependent pAOX1
344 promoter for inducible expression, so we selected methanol as the initial carbon source. We
345 then examined the impact of the source of nitrogen and buffer pH on titer. We conducted a full-
346 factorial DOE using identical concentrations as those used to create a medium for accumulating
347 biomass. The resulting model was visualized by ranking combinations of sources of nitrogen
348 and buffer (**Figure 3A**). The effects showed no interaction between these two factors. Urea was
349 again identified as the preferred source of nitrogen while higher pH values led to improved
350 secreted P[8] productivity. Unlike biomass accumulation, this pH dependence was observed
351 across both nitrogen sources.

352 We next applied the same DOE to identify important concentration-dependent
353 interactions that impact the production of P[8]. Unsurprisingly, the concentration of methanol
354 was the most important factor, with possible minor effects from other components (**Figure 3B**).
355 We decided to screen further a 20-fold range in methanol concentrations using two formulations
356 for remaining media components—the one determined for optimal cell growth (DM1) and the
357 optimal base media formulation predicted by the quadratic model here (2x YNB, 1 g/L urea, 4
358 g/L potassium phosphate adjusted to a pH of 6.5). We found that production was relatively

359 insensitive for concentrations of methanol ranging from 1-4 v/v%, with an optimum around 2%
360 (**Figure 3C**). We postulated that the rapid decline in productivity observed in these milliliter-
361 scale cultures using concentrations >6 v/v% methanol was likely due to excess formation of
362 toxic metabolic byproducts such as formaldehyde and hydrogen peroxide (Wakayama et al.,
363 2016). Interestingly, the predicted optimal medium from this set of studies outperformed the
364 medium we determined for accumulating biomass, suggesting that certain components of the
365 basal medium may benefit protein expression more than cellular growth and underscores the
366 value of optimizing media for specific phenotypes of interest. Based on these data in total, we
367 defined a basal medium for production including 2x YNB, 2 v/v% methanol, 1 g/L urea, and 4
368 g/L potassium phosphate buffer adjusted to a pH of 6.5 (DM2_dev0).
369 Next, we examined which supplements could improve the performance of DM2_dev0.
370 We added three chemical chaperones (TUDCA, sodium deoxycholate monohydrate (SDM), and
371 valproic acid) (Kuryatov, Mukherjee, & Lindstrom, 2013; Uppala, Gani, & Ramaiah, 2017), two
372 antioxidants (reduced glutathione (GSH) and N-acetyl cysteine (NAC)), and the chelator, K-
373 ETDA, to the list of 16 supplements included in our original screen defined for biomass
374 accumulation. Concentrations for these components were chosen based on product
375 specifications, literature data, and prior experience (Supporting Information). Many of the 22
376 supplements screened improved production of P[8] (**Figure 3D**). The top four ranking
377 supplements comprised surfactants or lipids, which could modulate membrane fluidity and lipid
378 metabolism (Butler, Huzel, Barnab, Gray, & Bajno, 1999; Degreif, Cucu, Budin, Thiel, & Bertl,
379 2019; Ritacco, Frank V; Yongqi Wu, 2018).
380 We then screened combinations of lipid supplements and surfactants to identify potential
381 synergistic effects. We ranked the individual supplements and their combinations (**Figure 3E**)
382 according to the measured titers of P[8]. We found that the addition of a cholesterol-rich
383 supplement yielded the highest secreted titers of P[8] (~50% improvement compared with
384 supplement-free condition in initial screens). Interestingly, a synthetic cholesterol supplement

385 alone did not substantially improve performance, suggesting the benefit results from a
386 combination of fatty acids and surfactant components in the supplement (**Supporting**
387 **Information**). This conclusion is consistent with similar improvements observed from other
388 supplements, such as linoleic acid-oleic acid-albumin (**Figure 3D**).

389 Since no other synergistic effects were observed in the combination screen, we
390 assessed the dependence of titer on the concentration of the cholesterol-containing supplement
391 identified (**Figure 3F**). Similar to our observations with cellular YNB used in the outgrowth
392 media, we found that concentrations of the supplement as low as 0.2 v/v% were beneficial for
393 protein expression, but that production was relatively insensitive to concentration (**Figures 3F,**
394 **3G**). We then directly compared the supplemented medium to the original composition; the new
395 supplemented media provided a 25% improvement in titer ($p = 0.0006$, one-tailed Welch's T
396 test). This new formulation with 1x cholesterol supplement, which we named DM2_dev1, was
397 the result of one cycle of optimization using our method.

398 Components of the cholesterol supplement included fatty acids, cholesterol, and
399 cyclodextrin, which are all are known to modulate membrane fluidity, a key parameter in vesicle
400 trafficking (Cooper, 1978; Degreif et al., 2019; Mahammad & Parmryd, 2015). We reasoned that
401 the addition of this supplement could therefore have synergistic effects with other supplements,
402 but did not find any further supplementation that improved P[8] titers within our original screen
403 (**Figure 3H**). We, therefore, considered if there could be additional classes of beneficial
404 supplements, absent from the original screen. Previous experiments demonstrated that P[8]
405 productivity is highly sensitive to methanol concentration (**Figure 3C**), so we wondered whether
406 further modulation of central carbon metabolism could yield additional productivity gains.

407 Modification of central carbon metabolism is best accomplished by feeding cells
408 alternative carbon sources, either entirely or as co-feeding substrates. Four co-fed substrates
409 have previously been shown to be non-repressive of pAOX1: sorbitol, mannitol, trehalose, and
410 alanine (Inan & Meagher, 2001). These substrates can be co-utilized with methanol without

411 repressing the pAOX1 promoter, which controls expression of P[8]. We hypothesized that the
412 introduction of supplemental carbon sources could enable further optimization of central carbon
413 metabolism. We screened co-fed substrates individually and in 1:1 combinations at a total
414 concentration of 20 g/L (a concentration similar to the optimal fructose and methanol
415 concentrations observed in previous carbon source optimizations) (**Figure 2E,3C**). Sorbitol co-
416 feeding had the most beneficial effect, resulting in a ~80% increase in P[8] titer (**Figure 3I**).
417 Mannitol supplementation was also beneficial (~70% increase), while alanine and trehalose co-
418 feeding were detrimental to productivity. While co-feeding carbon sources led to increased
419 biomass yield during production, these differences did not account for the improved titer, as
420 improvements in specific productivity (q_p) of ~60% and ~45% were also observed for the sorbitol
421 and mannitol co-fed conditions, respectively (**Supporting Information**). Based on these data,
422 we chose to include sorbitol as a supplemental carbon source for further study.
423 The addition of a supplemental carbon source could significantly impact central carbon
424 metabolism. We, therefore, wondered how the inclusion of sorbitol might impact the optimal
425 carbon feeding strategy. Examining total carbon source concentrations from 20 – 70 g/L, we
426 compared the performance of cultures co-fed with sorbitol:methanol ratios of 3:1, 1:1, and 1:3 to
427 a methanol-only control (**Figure 3J**). All co-fed conditions outperformed the methanol-only
428 control, suggesting that the presence of sorbitol is highly beneficial for producing P[8]. The titer
429 was relatively insensitive to sorbitol:methanol ratios and carbon concentrations. Based on the
430 data, we decided to use 2 v/v% methanol and 20 g/L of sorbitol for the final sorbitol-
431 supplemented media named DM2.
432 Finally, we compared the P[8] titer obtained using DM2_dev0, DM2_dev1, and DM2 to
433 other common production media for *P. pastoris*: BMMY and RDM. We found that DM2 led to a
434 ~2x improvement in P[8] titers, relative to BMMY and RDM, up to 97 $\pm$ 2 mg/L.
435
436 **Discussion**

437 Here we have implemented a novel and broadly applicable approach for media
438 development that relies on rapid, automated construction of diverse media from defined
439 modules of components. We demonstrated the utility of this approach by developing two new
440 media for two phenotypes of interest in the heterologous production of proteins by yeast,
441 namely biomass accumulation and secreted production. We systematically identified and
442 optimized the concentration of media components important to each phenotype of interest.
443 Importantly, defining these new formulations of media did not require advanced analytical
444 capabilities and required minimal experimental time to assess more than 360 total formulations
445 during two to three rounds of optimization for each.
446 Our optimized formulations affirmed the importance of lipid-related components for
447 maximizing titers in *Pichia pastoris* cultivations. The importance of optimizing membrane fluidity
448 or lipid metabolism has been well established in CHO and appears to be key to optimizing
449 heterologous protein secretion in *P. pastoris* cultivation as well (Clincke et al., n.d.; Ritacco et
450 al., 2018; Zhang, Wang, & Liu, 2013).
451 Modular media blending has four advantages over existing methods. First, the use of
452 common stock solutions and supplements to formulate media reduces initial labor required for
453 new experiments or optimizations ~15 minutes per experiment, making parallel testing of
454 multiple hypotheses efficient and requires less resources overall. Here, we created 30 stock
455 solutions, and evaluated >360 unique media compositions, without manual preparation of
456 individual media or extensive blending calculations or planning. Most of these solutions could be
457 readily reused in future experiments to optimize for new phenotypes of interest. Second, our
458 method requires minimal knowledge of the host organism *a priori* and could, in principle, be
459 applied to any measurable phenotype of interest. We anticipate that this method could be used
460 to optimize other phenotypes of interest, such as glycosylation profiles. Third, our method
461 provides certain practical advantages, including minimal requirements for analytical
462 characterization and rapid identification of component interactions that lead to solubility

463 challenges. These traits make it possible to learn about formulations that may lead to extensive
464 precipitates like those encountered with our rich defined medium formulation (**Figure 4A**).
465 Finally, modularly constructed media, such as DM2, can be ~70% pure water with low
466 osmolarity, leaving volumetric and osmotic space for future modifications to accommodate new
467 or related phenotypes of interest (**Figure 4B**).

468 We also acknowledge certain limitations in the present study that may be addressed in
469 future work. First, while modular media development identifies components key to the
470 optimization of the phenotype of interest, additional media optimization effort may be necessary
471 to translate these learning in batch cultivations to scaled-up fed-batch or perfusion operation,
472 where additional variables such as supplemental feed composition and feeding schedule must
473 also be considered. In principle, modular media construction could be applied to high-
474 throughput scale-down cultivation models, such as Ambr250s. Second, our approach for
475 optimization relies on greedy algorithms tailored to create a new media for a single phenotype of
476 interest; however, given the vast explorable media space it is possible to find a local optimum.
477 Further metabolic or -omic modeling techniques could be employed to guide broader exploration
478 of media space, co-optimize multiple phenotypes, or facilitate biologically informed optimization,
479 albeit with more complex experimental and computational requirements (Matthews, Kuo, Love,
480 & Love, 2017b; Mohmad-Saberi et al., 2013). Third, our current method used commercially
481 available supplements, but in practice, beneficial supplements could be simplified by using
482 individual components, to facilitate more biological inferences and aid development of improved
483 host-specific supplements. Finally, initial screens to identify beneficial supplements rely on
484 reasonable choices of initial concentrations for screening. These currently require prior
485 knowledge from the literature or commercial sources; with further use in the community of the
486 Openblend approach, it is possible additional sharing of knowledge could help inform further
487 developments.

The improved speed and accessibility of in-depth media development experiments enabled by modular media construction could help improve expression of many classes of proteins in laboratories and discovery centers that have not traditionally had access to such capabilities. Since many lead candidates for new therapeutic proteins begin in small biotech firms and academic labs, early-stage improvements in productivity could help advance more proteins towards the clinic simply by facilitating access to larger quantities of proteins for initial research and non-clinical studies. In more established companies, the ability to make rapid improvements to existing media may enable faster product development timelines and could reduce manufacturing costs overall. Rapid identification and optimization of sensitive media components could also enable easier adoption of a range of industrially relevant alternative hosts, resulting in further manufacturing flexibility and potentially cost savings (Coleman, 2020).

Acknowledgements

The authors acknowledge Danielle Camp for program coordination. This work was funded by the Bill & Melinda Gates Foundation (Investment ID INV-002740). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Bill & Melinda Gates Foundation.

A.M.B., I.R.G., K.R.L., and J.C.L. conceived and planned experiments. A.M.B. conducted media development experiments. S.R.A. developed and maintained the RP-UPLC assay. A.M.B. performed analytical characterization. A.M.B., K.R.L., and J.C.L. wrote the manuscript. J.C.L. and K.R.L. designed the experimental strategy and supervised analysis. All authors reviewed the manuscript.

References

Butler, M., Huzel, N., Barnab, N., Gray, T., & Bajno, L. (1999). Linoleic acid improves the robustness of cells in agitated cultures, 27–36.

Clincke, M.-F., Guedon, E., Yen, F. T., Ea, L., Universite, N., Ogier, V., … Goergen, J.-L. (n.d.). Effect of Surfactant Pluronic F-68 on CHO Cell Growth , Metabolism , Production , and Glycosylation of Human Recombinant IFN- c in Mild Operating Conditions, 181–190. https://doi.org/10.1002/btpr.503

Coleman, E. (2020). Establishment of a Novel Pichia Pastoris Host Production Platform by.

Combs, G. F. (2012). *The vitamins: Fundamental aspects in nutrition and health*. San Diego: Elsevier Academic Press.

Cooper, R. A. (1978). Influence of Increased Membrane Cholesterol on Membrane Fluidity and Cell Function in Human Red Blood Cells. *Journal of Supramolecular Structure*.

Crowell, L. E., Crowell, L. E., Raymond, A., St, H. E., Engineering, C., Doyle, P. S., & Crowell, L. E. (2020). Accelerating process development for biologics on an automated , pharmacy-scale manufacturing system by by.

Dalvie, N. C., Brady, J. R., Crowell, L. E., Tracey, M. K., Biedermann, A. M., Kaur, K., … Love, J. C. (2020). Molecular engineering improves antigen quality and enables integrated manufacturing of a trivalent subunit vaccine candidate for rotavirus. *BioRx*, 1–51.

Degreif, D., Cucu, B., Budin, I., Thiel, G., & Bertl, A. (2019). Lipid determinants of endocytosis and exocytosis in budding yeast. *BBA - Molecular and Cell Biology of Lipids*, *1864*(7), 1005–1016. https://doi.org/10.1016/j.bbalip.2019.03.007

FSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP). (2011). Scientific Opinion on safety and efficacy of choline chloride as a feed additive for all animal species. *EFSA Journal*, *9*(9), 2353.

Gagnon, M., Hiller, G., Luan, Y. T., Kittredge, A., Defelice, J., & Drapeau, D. (2011). High-End pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO Fed-batch cultures. *Biotechnology and Bioengineering*, *108*(6), 1328–1337. https://doi.org/10.1002/bit.23072

Galbraith, S. C., Bhatia, H., Liu, H., & Yoon, S. (2018). Media formulation optimization: current and future opportunities. *Current Opinion in Chemical Engineering*, *22*, 42–47. https://doi.org/10.1016/j.coche.2018.08.004

Guo, C., Huang, Y., Zheng, H., Tang, L., He, J., Xiang, L., … Jiang, H. (2012). Secretion and activity of antimicrobial peptide cecropin D expressed in Pichia pastoris. *Experimental and Therapeutic Medicine*, 1063–1068. https://doi.org/10.3892/etm.2012.719

Inan, M., & Meagher, M. M. (2001). Non-repressing carbon sources for alcohol oxidase (AOX1) Promoter of Pichia pastoris, *92*(6), 585–589.

Jordan, M., Voisard, D., Berthoud, A., & Tercier, L. (2013). Cell culture medium improvement by rigorous shuffling of components using media blending, 31–40. https://doi.org/10.1007/s10616-012-9462-1

Kennedy, M., & Krouse, D. (1999). Strategies for improving fermentation medium performance: a review. *Journal of Industrial Microbiology and Biotechnology*, (23), 456–475.

Kuryatov, A., Mukherjee, J., & Lindstrom, J. (2013). Chemical Chaperones Exceed the Chaperone Effects of RIC-3 in Promoting Assembly of Functional a 7 AChRs. *PLoS ONE*, *8*(4), 1–11. https://doi.org/10.1371/journal.pone.0062246

Loebrich, S., Clark, E., Ladd, K., Takahashi, S., Brousseau, A., Kitchener, S., … Ryll, T. (2019). Comprehensive manipulation of glycosylation profiles across development scales. *MAbs*, *11*(2), 335–349. https://doi.org/10.1080/19420862.2018.1527665

Love, J. C., Love, K. R., & Barone, P. W. (2012). Enabling global access to high-quality biopharmaceuticals. *Current Opinion in Chemical Engineering*, *2*(4), 383–390. https://doi.org/10.1016/j.coche.2013.09.002

Lu, T. L., Pugach, O., Somerville, R., Rosenberg, S. A., Kochenderfer, J. N., Better, M., & Feldman, S. A. (2016). A Rapid Cell Expansion Process for Production of Engineered Autologous CAR-T Cell Therapies. *Human Gene Therapy Methods*, *27*(6), 209–219. https://doi.org/10.1089/hgtb.2016.120

Mahammad, S., & Parmryd, I. (2015). Cholesterol depletion using methyl-β-cyclodextrin. In *Methods in membrane lipids* (pp. 91–102).

Matthews, C. B., Kuo, A., Love, K. R., & Love, J. C. (2017a). Development of a general defined medium for Pichia pastoris, (July), 103–113. https://doi.org/10.1002/bit.26440

Matthews, C. B., Kuo, A., Love, K. R., & Love, J. C. (2017b). Development of a general defined medium for Pichia pastoris, (July), 103–113. https://doi.org/10.1002/bit.26440

McGillicuddy, N., Floris, P., Albrecht, S., & Bones, J. (2018). Examining the sources of variability in cell culture media used for biopharmaceutical production. *Biotechnology Letters*, *40*(1), 5–21. https://doi.org/10.1007/s10529-017-2437-8

Mohmad-Saberi, S. E., Hashim, Y. Z. H. Y., Mel, M., Amid, A., Ahmad-Raus, R., & Packeer-Mohamed, V. (2013). Metabolomics profiling of extracellular metabolites in CHO-K1 cells cultured in different types of growth media. *Cytotechnology*, *65*(4), 577–586. https://doi.org/10.1007/s10616-012-9508-4

Moser, J. W., Prielhofer, R., Gerner, S. M., Graf, A. B., Wilson, I. B. H., Mattanovich, D., & Dragosits, M. (2017). Implications of evolutionary engineering for growth and recombinant protein production in methanol - based growth media in the yeast Pichia pastoris. *Microbial Cell Factories*, 1–16. https://doi.org/10.1186/s12934-017-0661-5

Pereira, S., Kildegaard, H. F., & Andersen, M. R. (2018). Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients. *Biotechnology Journal*, *13*(3), 1–13. https://doi.org/10.1002/biot.201700499

Ritacco, Frank V; Yongqi Wu, A. K. (2018). Cell Culture Media for Recombinant Protein Expression in Chinese Hamster Ovary (CHO) Cells : History , Key Components , and Optimization Strategies. https://doi.org/10.1002/btpr.2706

Ritacco, F. V., Wu, Y., & Khetan, A. (2018). Cell culture media for recombinant protein expression in Chinese hamster ovary (CHO) cells: History, key components, and optimization strategies. *Biotechnology Progress*, *34*(6), 1407–1426. https://doi.org/10.1002/btpr.2706

Rodrigues, M. E., Costa, A. R., Henriques, M., Azeredo, J., & Oliveira, R. (2012). Comparison of commercial serum-free media for CHO-K1 cell growth and monoclonal antibody production. *International Journal of Pharmaceutics*, *437*(1–2), 303–305. https://doi.org/10.1016/j.ijpharm.2012.08.002

Rouiller, Y., Périlleux, A., Collet, N., Jordan, M., Stettler, M., & Broly, H. (2013). A high-throughput media design approach for high performance mammalian fed-batch cultures, (June), 501–511.

Schnellbaecher, A., Binder, D., Bellmaine, S., & Zimmer, A. (2019). Vitamins in cell culture media: Stability and stabilization strategies. *Biotechnology and Bioengineering*, *116*(6), 1537–1555. https://doi.org/10.1002/bit.26942

Sorkun, M. C., Khetan, A., & Er, S. (2019). AqSolDB , a curated reference set of aqueous solubility and 2D descriptors for a diverse set of compounds, *2019*(July), 1–8. https://doi.org/10.1038/s41597-019-0151-1

Uppala, J. K., Gani, A. R., & Ramaiah, K. V. A. (2017). Chemical chaperone , TUDCA unlike PBA , mitigates protein aggregation efficiently and resists ER and non-ER stress induced HepG2 cell death. *Scientific*, *1*(January), 1–13. https://doi.org/10.1038/s41598-017-03940-1

Villadsen, J. (2015). Redox Balances and Consistency Check of Experiments. In *Fundamental Bioengineering* (pp. 17–38).

Wakayama, K., Yamaguchi, S., Takeuchi, A., Mizumura, T., Ozawa, S., Tomizuka, N., …

Nakagawa, T. (2016). Regulation of intracellular formaldehyde toxicity during methanol metabolism of the methylotrophic yeast Pichia methanolica. *Journal of Bioscience and Bioengineering*, *122*(5), 545–549. https://doi.org/10.1016/j.jbiosc.2016.03.022

Yamamoto, T., & Ishihara, K. (n.d.). Stability of Glutathione in Solution. *Developments in Food Engineering*, 209–211.

Zakharova, E. M., & Minashina, I. K. (2015). Review of Multidimensional Optimization Methods. *Mathematical Models, Computational Methods*, *60*(6), 625–636. https://doi.org/10.1134/S1064226915060194

Zhang, H., Wang, H., & Liu, M. (2013). Rational development of a serum-free medium and fed-batch process for a GS-CHO cell line expressing recombinant antibody, 363–378. https://doi.org/10.1007/s10616-012-9488-4

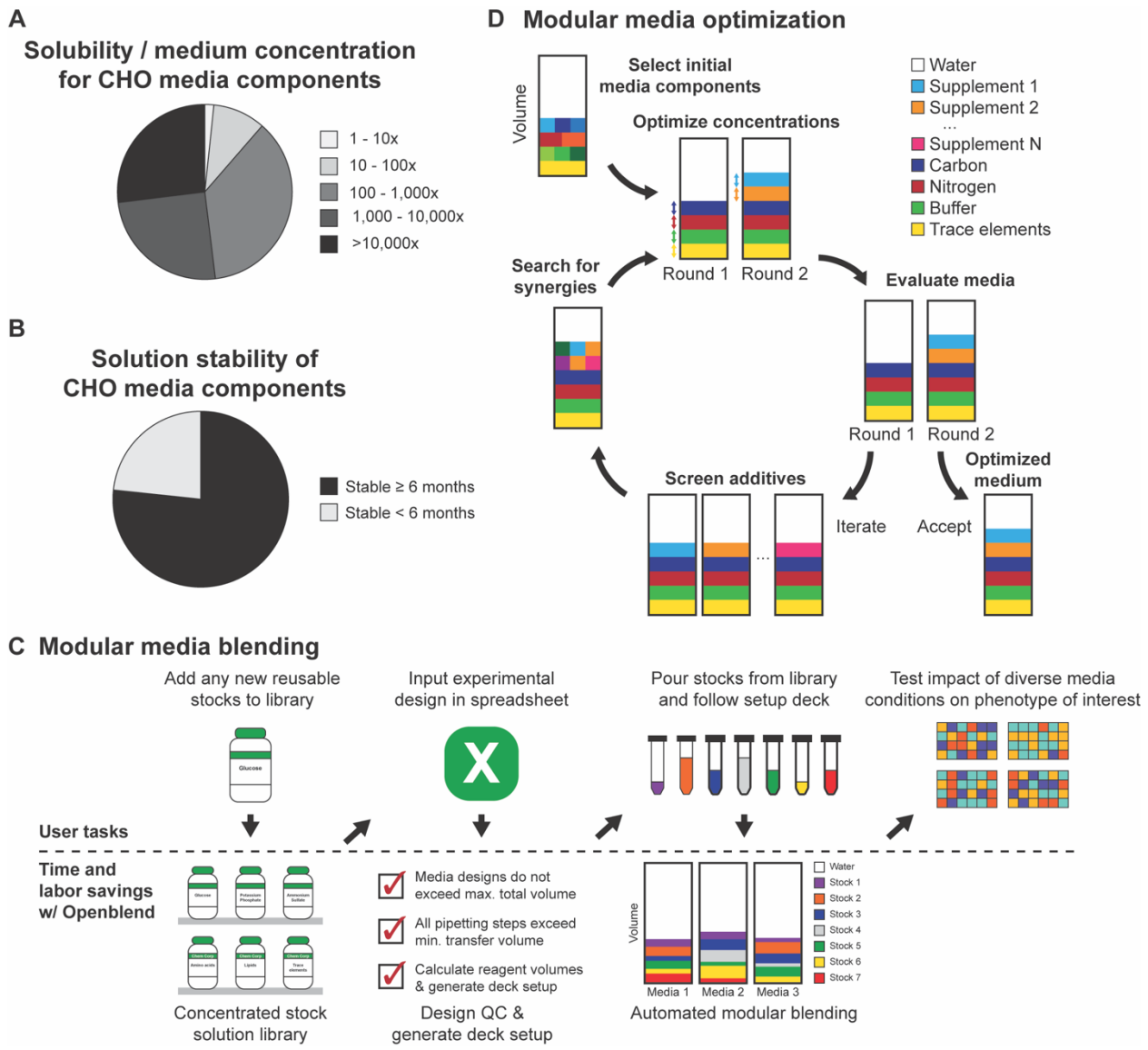


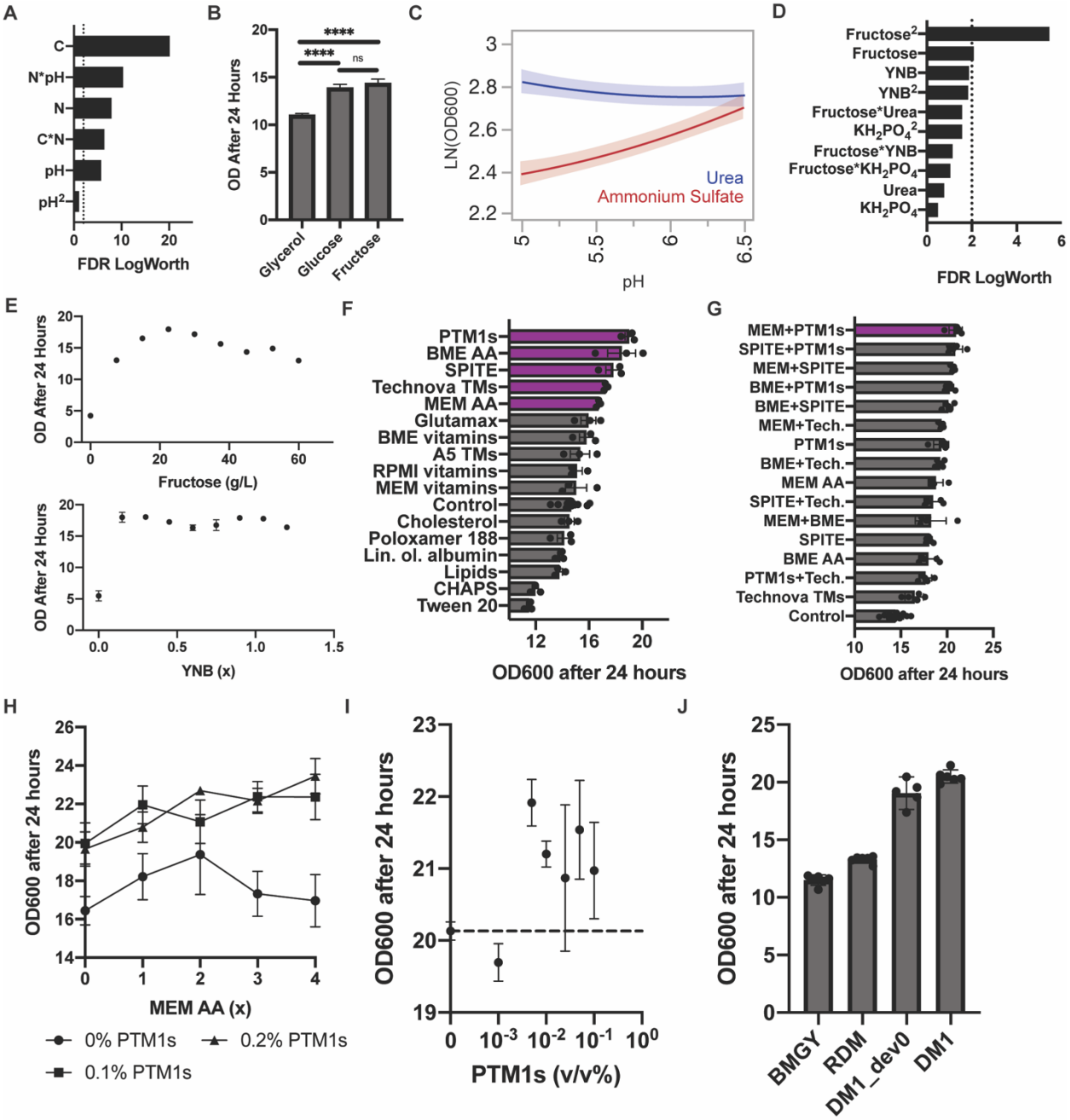
Figure 1. Modular media development can be broadly applicable, easily applied, and systematically executed to improve measurable phenotypes of interest.

A) Estimate of the ratios of component solubility to their concentrations in medium demonstrates that most components are soluble at >10x their concentration in the CHO medium, eRDF. B) With the exception of some classes of medium components, such as vitamins, most media components can be formulated into solutions that remain stable for >6 months under proper storage conditions. C) Overview of time, labor and planning saved by using Openblend to automate modular media construction. D) Overview of a modular media optimization approach,

636 which can be used to build an optimized medium for any measurable phenotype of interest

637 systematically.

638



639

640 **Figure 2**. Modular development of a new biomass accumulation media for *P. pastoris*

641 A) Significance of carbon (fructose, glucose, glycerol), nitrogen (urea and ammonium sulfate),

642 and pH choice (5, 5.75, 6.5) in a least square regression model fitted to a full factorial DOE. B)

643 Fructose and glucose were found to result in significantly higher biomass accumulation after 24

hours of outgrowth than glycerol. C) Ammonium sulfate was found to be more pH sensitive than urea, as shown by the JMP sensitivity profiles during fructose feeding. D) Significance of terms in a least square regression model fitted to a full factorial DOE over fructose, urea, potassium phosphate, and YNB concentrations. E) 1-FAAT optimization of fructose and YNB concentration finds optimal outgrowth performance at a fructose concentration of 22.5 g/L and relative insensitivity over a wide range of YNB concentrations (0.15 to 1.2x). F) A media supplementation screen identified 5 beneficial supplements, related to trace element and amino acid supplementation. G) Further screening of beneficial supplement combinations identified synergistic amino acid and trace metal supplementation strategies. H) Comparison of the effect of MEM amino acid concentration on biomass accumulation at different PTM1 salts concentrations. I) Effect of the concentration of PTM1 salts on biomass accumulation in DM1_dev0 medium supplemented with 1x MEM AA. J) Head-to-head comparison of 4 v/v% glycerol BMGY, 4 v/v% glycerol rich defined medium, the initial defined biomass accumulation media (DM1_dev0), and the final biomass accumulation medium obtained after a full optimization cycle (DM1), demonstrates that DM1 leads to superior biomass accumulation.

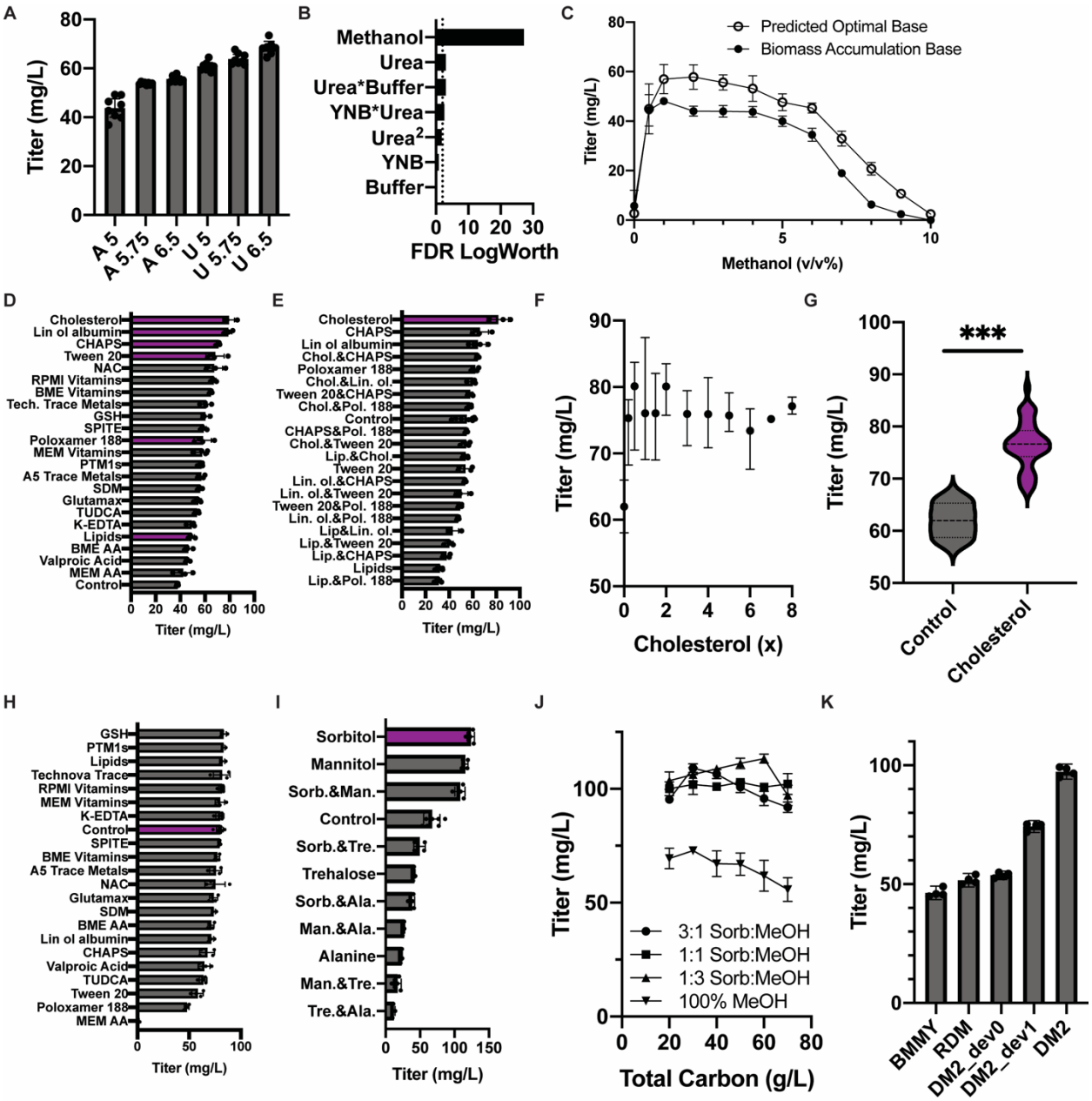


Figure 3. Modular development of a media for heterologous protein production in *P. pastoris* A) Initial full-factorial screen of nitrogen source choice and buffer pH demonstrates that urea is preferred over ammonium sulfates and high buffer pH is preferred over lower values. B) A full-factorial concentration optimization identified methanol as the most concentration dependent variable. Other components in the base media were predicted to affect productivity with much lower levels of significance. C) Evaluation of the effect of methanol concentration on P[8] titer, using two different base media (urea, buffer, and YNB concentrations): the biomass

accumulation base medium and the optimal base media composition predicted by our concentration DOE. D) Ranking of supplements according to their effect on P[8] titer. Supplements related to membrane fluidity or lipid metabolism ranked highly. E) Evaluation of combinations of lipid and surfactant supplements confirmed that cholesterol supplementation leads to the greatest improvement in P[8] titer. F) Concentration optimization of cholesterol demonstrated low concentration dependence, with similar performance observed over a 40-fold range (0.2-8x). G) Comparing cholesterol-free and cholesterol-supplemented cultures fed at various concentrations demonstrates that cholesterol supplementation results in a significant ~25% improvement in P[8] titers ($p<0.001$). H) No significantly beneficial supplements were observed when repeating the supplementation screen. I) Screening supplementation of 20 g/L of co-fed substrates individually or in 1:1 combinations by mass identified sorbitol supplementation as highly beneficial to P[8] titer. J) Examination of the effect of co-feed ratio and total carbon concentration on titer in DM2_dev1 supplemented media. K) Comparison of P[8] titer obtained with DM2 to previous iterations and other common *P. pastoris* media demonstrates a ~2x improvement in P[8] titer, relative to 1 v/v% methanol RDM and 1 v/v% methanol BMMY.

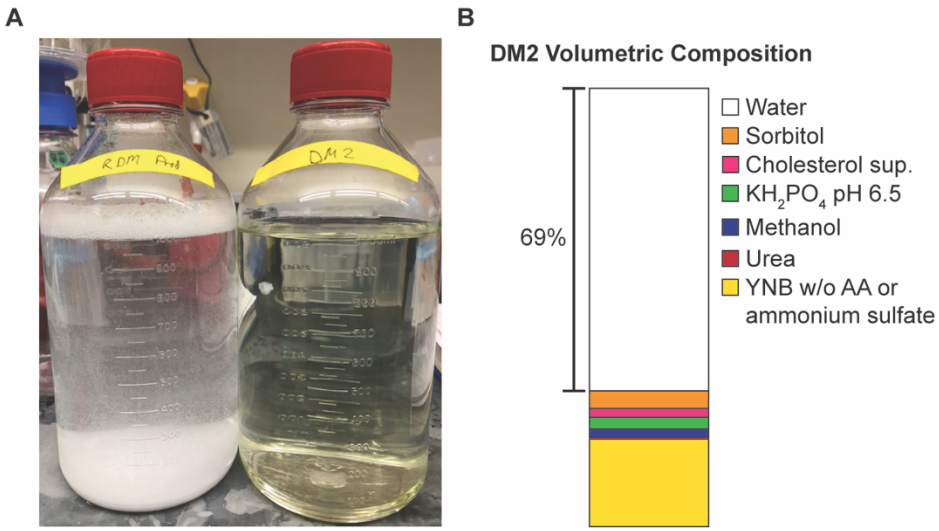


Figure 4. Comparison of DM2 to rich define medium.

A) Comparison of precipitate formation during construction of RDM (left) and DM2 (right) media. Adjusting the pH of RDM to 6.5 results in significant formation of white precipitate. No precipitate formation is observed in DM2. B) Relative volumes of stock solutions and pure water needed to construct DM2. Pure water addition accounts for 69% of DM2 volume, demonstrating that there is substantial room for further supplement exploration and development. When separated into simple stock solutions, DM2 can be 3x concentrated.