

# Microscale to Manufacturing Scale-up of Cell-Free Cytokine Production—A New Approach for Shortening Protein Production Development Timelines

James F. Zawada,<sup>1</sup> Gang Yin,<sup>1</sup> Alexander R. Steiner,<sup>1</sup> Junhao Yang,<sup>1</sup> Alpana Naresh,<sup>1</sup> Sushmita M. Roy,<sup>1</sup> Daniel S. Gold,<sup>2</sup> Henry G. Heinsohn,<sup>1</sup> Christopher J. Murray<sup>1</sup>

<sup>1</sup>Sutro Biopharma, Inc. 310 Utah Ave. Suite 150, South San Francisco, California 94080; telephone: 650-888-6329; fax: 650-872-8924; e-mail: cmurray@sutrobio.com

<sup>2</sup>Rosebud Biotech, 3223 Santiago Street, San Francisco, California 94116

Received 27 November 2010; revision received 29 January 2011; accepted 7 February 2011

Published online 17 February 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bit.23103

**ABSTRACT:** Engineering robust protein production and purification of correctly folded biotherapeutic proteins in cell-based systems is often challenging due to the requirements for maintaining complex cellular networks for cell viability and the need to develop associated downstream processes that reproducibly yield biopharmaceutical products with high product quality. Here, we present an alternative *Escherichia coli*-based open cell-free synthesis (OCFS) system that is optimized for predictable high-yield protein synthesis and folding at any scale with straightforward downstream purification processes. We describe how the linear scalability of OCFS allows rapid process optimization of parameters affecting extract activation, gene sequence optimization, and redox folding conditions for disulfide bond formation at microliter scales. Efficient and predictable high-level protein production can then be achieved using batch processes in standard bioreactors. We show how a fully bioactive protein produced by OCFS from optimized frozen extract can be purified directly using a streamlined purification process that yields a biologically active cytokine, human granulocyte-macrophage colony-stimulating factor, produced at titers of 700 mg/L in 10 h. These results represent a milestone for in vitro protein synthesis, with potential for the cGMP production of disulfide-bonded biotherapeutic proteins.

Biotechnol. Bioeng. 2011;108: 1570–1578.

© 2011 Wiley Periodicals, Inc.

**KEYWORDS:** cell-free protein synthesis; disulfide bond formation; transcription–translation

## Introduction

Cell-free protein synthesis systems have distinct advantages over traditional in vivo methods for protein production (Endo and Sawasaki, 2006; Jermutus et al., 1998). The absence of a requirement to maintain cell viability allows optimization of the protein synthetic capacity of the cell-free extract to produce a single protein. The absence of a cell wall allows for addition of non-natural factors to the open system in order to manipulate transcription, translation, and folding to provide precise modulation of the protein expression process (Jewett and Swartz, 2004b). Cell-free systems derived from *Escherichia coli* have a long history as experimental tools for exploring the molecular biology of transcription and translation (Jewett et al., 2002; Noren et al., 1989; Spirin et al., 1988; Zubay, 1973) and have been utilized as a basis for developing in vitro protein production systems at small scales.

Cell-free protein production at the multigram and kilogram scale, an essential starting requirement for biotherapeutic production processes, has been hampered by the lack of scalable systems amenable to standard bioreactor configurations at large scale. Voloshin and Swartz (2008) reported production of 385 mg/L of soluble protein corresponding to a 3-disulfide bonded protein IGF-1 to 1-L scale. However, the fraction of correctly folded and functional protein was not reported, and different yields were reported using different reaction vessel designs. Consistent and robust performance at a wide range of scales is important for using a system as a general production technology and the time and cost of biologic drug discovery and production process development can be substantially reduced with systems that scale in a predictable and efficient manner from high-throughput screening to manufacturing.

Correspondence to: C.J. Murray

Contract grant sponsor: National Science Foundation

Contract grant number: NSF0420208; NSF0522337

Contract grant sponsor: National Science Foundation

Additional Supporting Information may be found in the online version of this article.

The industrial scale application of cell-free protein synthesis for the production of recombinant proteins requires efficient and reproducible preparation of large volumes of active cell extracts (Liu et al., 2005), stable and cost-effective energy generating systems (Jewett and Swartz, 2004a), and the maintenance of stable pools of precursor molecules required for high-level protein synthesis. For example, the overall protein synthetic capacity of *E. coli* extracts has been increased by genetically engineering pathways responsible for recycling key amino acid precursor molecules, leading to the development of the *E. coli* strain KGK10 (Calhoun and Swartz, 2006; Knapp et al., 2007; Michel-Reydellet et al., 2004).

Here, we extend the *E. coli*-based system developed by Swartz and coworkers to show how scalable open cell-free synthesis (OCFS) can be optimized for high-level production of a multi-disulfide-bonded protein, granulocyte-macrophage colony-stimulating factor (rhGM-CSF), through detailed analysis and refinement of several process elements including: extract preparation and storage, optimization of synthetic gene sequences and redox parameters, process scale-up, and downstream product purification steps. We show that after full optimization, the OCFS system can produce active rhGM-CSF, a four-helix bundle human cytokine with two disulfide bonds required for functional activity. The system is linearly scalable and can achieve yields of rhGM-CSF of 700 mg/L in 10 h over  $\sim 10^6$  range in volumes to 100 L. This linear scalability allows for rapid analysis of multiple parameters affecting transcription, protein synthesis, and folding in 24- or 96-well plates for high-throughput optimization of product and process parameters. Finally, the absence of a cell-lysis step after protein production, allows rapid and consistent protein purification optimization.

## Materials and Methods

### Cell Extract Prepared From Strain KGK10 at 200 L Scale

Cell extracts were prepared using a previously engineered K-12-derived *E. coli* strain KGK10 [A19  $\Delta endA$   $\Delta tonA$   $\Delta speA$   $\Delta tnaA$  met<sup>+</sup>  $\Delta sdaA$   $\Delta sdaB$   $\Delta gshA$   $\Delta gor$  *TrxB\_HA* tag; (Knapp et al., 2007)]. KGK10 was cultured to mid-log phase [OD<sub>595</sub>  $\sim$  45 OD,  $\sim$  140 g/L of cell wet weight (CWW)], using glucose and amino acid fed-batch fermentation at maximal growth rate of 0.7 h<sup>-1</sup>, essentially as described by (Zawada and Swartz, 2005), except that the glucose feed rate was increased manually to ensure excess glucose was present at the time of harvest. Fermentations were performed in a 200 L Sartorius D200 bioreactor equipped with pH and temperature control, four baffles, three six-bladed impellers, sampling tube, air sparger, pO<sub>2</sub>-probe, and exhaust gas analysis. Cells were pelleted by centrifugation at  $>14,000g$  for 45 min at 4°C two times using tubular bowl centrifuges (Model Z101; Carl Padberg, Schnell Zentrifugenbau GmbH, Lahr, Germany) or a disc-stack continuous centrifuge

(Model CSC6 with hydrostop; GEA Westfalia GmbH, Oelde, Germany) with a maximum bowl speed of 12,000 rpm and a feed flow rate of 2.0–2.5 L/min, then resuspended and repelleted two times with S30 buffer and stored at  $-80^\circ\text{C}$ .

The thawed cell-paste was suspended in 2 mL of S30 buffer per g CWW, cells were lysed in a homogenizer (EmulsiFlex-C55A; Avestin, Mannheim, DE) at 20,000 psi and chilled to 0–4°C using a heat exchanger on the homogenizer outlet. Cell debris and insoluble components were removed by centrifugation at 12,000g for 45 min at 4°C two times. A typical 200 L fermentation yielded  $>1.1$  L clarified extract/kg CWW with a total protein concentration of 20–25 g/L. For optimization of extract activity (“run-off reaction”), clarified extract samples (1 mL) were collected at various pre-incubation times at 25, 30, and 37°C, centrifuged at  $>15,000g$  to remove particulates, and then the supernatant was flash frozen in liquid nitrogen for subsequent analysis of protein synthesis activity, measured by rhGM-CSF production as described below. Once optimal pre-incubation conditions were identified, the extracts were activated at large scale and centrifuged at 14,000g for 35 min in tubular bowl centrifuges. The supernatant was frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  in polyethylene bags or, for long-term stability studies, in Freeze-Pak Bio-container bags (60 mL; Charter Medical Ltd, Winston-Salem, NC), or at  $-40^\circ\text{C}$  in cryostorage vials.

### Gene Expression Constructs

Expression of rhGM-CSF (GenBank Accession No. AAA52578.1) with an N-terminal methionine was under the control of the T7 promoter using added T7 RNA polymerase (RNAP) for transcription. Genes were designed using Biomax ProteoExpert (<https://ssl.biomax.de/ProteoExpert/index.jsp>) or DNA 2.0 GeneDesigner (<https://www.dna20.com/genedesigner2/>) algorithms (Welch et al. 2009) and were synthesized (DNA 2.0, Menlo Park, CA). Detailed methods for large-scale production of plasmid DNA and accessory proteins T7 RNAP and *E. coli* DsbC are described in Supplementary Methods.

### OCFS Protein Synthesis Reaction Conditions

Cell-free extracts were thawed to room temperature and incubated with 50  $\mu\text{M}$  iodoacetamide for 30 min, as previously described (Knapp et al. 2007). OCFS reactions were run at 30°C for up to 10 h containing 30% (v/v) iodoacetamide-treated extract with 8 mM magnesium glutamate, 10 mM ammonium glutamate, 130 mM potassium glutamate, 35 mM sodium pyruvate, 1.2 mM AMP, 0.86 mM each of GMP, UMP, and CMP, 2 mM amino acids (1 mM for tyrosine), 4 mM sodium oxalate, 1 mM putrescine, 1.5 mM spermidine, 15 mM potassium phosphate, 100 nM T7 RNAP, 2–10  $\mu\text{g/mL}$  plasmid DNA template, 1–10  $\mu\text{M}$  *E. coli* DsbC. Reduced (GSH) and oxidized (GSSG) glutathione were added to a total

concentration of  $\sim 5$  mM. The initial redox potential was calculated using the Nernst equation with  $E^0 = -205$  mV as the standard potential of the GSH/GSSG couple at  $30^\circ\text{C}$  and pH 7 (Wunderlich and Glockshuber, 1993).

Proteins produced at  $250\ \mu\text{L}$  scale or less were monitored by incorporation of  $1\text{-}[U\text{-}^{14}\text{C}]\text{-leucine}$  ( $300\ \mu\text{Ci}/\mu\text{mole}$ ; GE Life Sciences, Piscataway, NJ) by measuring the TCA-precipitable soluble and total protein in 24-well plates or Petri dishes without shaking (Voloshin and Swartz, 2005), or in 96-well plates at  $30\ \mu\text{L}$  scale (with shaking). Proteins were analyzed by reducing or non-reducing 12% SDS-PAGE gels with Sypro staining (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations and analyzed by autoradiography using a Storm 840 PhosphorImager.

At larger scales, concentrations of rhGM-CSF were determined using reversed phase HPLC (ZORBAX 300SB-C8  $5\ \mu\text{m}$ ,  $4.6 \times 250$  mm; Agilent, Santa Clara, CA) at  $40^\circ\text{C}$  with a flow rate of  $1.5\ \text{mL}/\text{min}$  in  $100\ \text{mM}$  triethylammonium acetate, pH 7.0 buffer. rhGM-CSF eluted 4 min into a 35–45.5%, acetonitrile gradient carried out over 7 min.

### OCFS Reaction Process Scale-up

OCFS reaction conditions were linearly scaled to 0.3 L in a 0.5 L stirred tank reactor (Biostat Q; B. Braun Biotech Inc., Bethlehem, PA), to the 4 L scale in a 10 L fermentor (ES10; B. Braun Biotech Inc.), and to the 100 L scale in a 200 L fermentor. Antifoam ( $1/12,000$  v/v of 31R1 Pluronic<sup>®</sup> surfactant; BASF, Florham Park, NJ) was added, the pH maintained at 7.0 and dissolved oxygen concentration  $[\text{O}_2] \sim 20\%$  by adjusting agitation speed and the  $[\text{O}_2]$  in the sparge gas at 1 vvm and an initial ratio of 90% air to 10%  $\text{O}_2$ . The 4 and 100 L scale reactions were pressurized at 5 psi to reduce foaming and  $1\ \mu\text{g}/\text{mL}$  ciprofloxacin was added to inhibit bacterial growth.

### Protein Purification

A portion of the rhGM-CSF was purified directly from the 100 L cell-free reaction as follows. Four liters of the 10 h cell-free reaction were diluted to 40 L with  $10\ \text{mM}$  sodium phosphate pH 6.5 and loaded onto 600 mL of DEAE Sepharose FastFlow resin (GE Life Sciences), washed with 10 L of  $10\ \text{mM}$  sodium phosphate pH 6.5, followed by gradient elution with  $10\ \text{mM}$  sodium phosphate, pH 6.5,  $0.2\ \text{M}$  NaCl over 30 column volumes. The DEAE purified product pool was concentrated 95-fold by tangential flow filtration with a 3 kD MW cutoff membrane (Pellicon 2 regenerated cellulose; Millipore, Billerica, MA). The concentrated product pool was then processed in multiple size exclusion chromatography runs using Sephacryl S-100 High-Resolution resin (GE Life Sciences) packed into two 1 L columns to give a combined column length of approximately 1 m. The overall recovery was 65%, and endotoxin and host cell protein levels were within

specifications (Supplementary Methods). The activity of rhGM-CSF was assayed by the dose-dependent stimulation of the proliferation of human TF-1 cells as previously described (Goerke and Swartz, 2008).

## Results

### Cell Fermentation and Extract Preparation

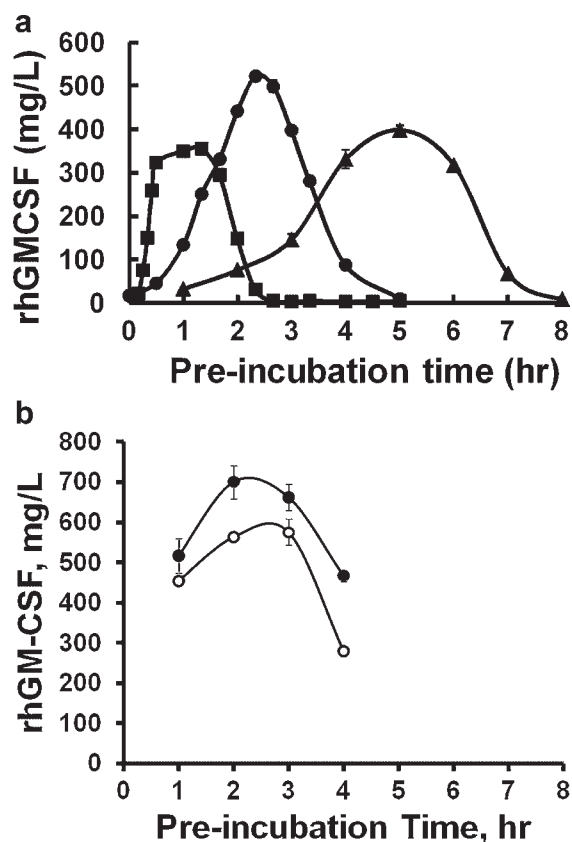
The OCFS bioproduction process starts with high-density cell culture with fast growth rates in order to optimize the yield of ribosomes that is directly correlated with cell growth rate (Zawada and Swartz, 2006). We modified the short protocol of Liu et al. (2005) for extract production using fermentation under conditions of excess glucose at harvest, rather than glucose-limited fermentation. Several extract process improvements included, introduction of high-throughput centrifugation using a disc-stack centrifuge capable of processing hundreds of liters of cells, increased dilution of the cell suspension prior to homogenization to improve subsequent lysate clarification, as well as removal of the requirement for a dialysis step following the optimized pre-incubation protocol, as described below.

### Optimized Extract Activation and Storage Stability

An essential step in the preparation of the extract is the pre-incubation prior to the initiation of T7-based transcription and translation of protein product, sometimes referred to as a translation “run-off” procedure or pre-incubation step (Zubay, 1973). Although the exact nature of the biochemical steps associated with cell-extract pre-incubation are not known in detail, we chose to investigate the effect of time and temperature of the pre-incubation step on the synthesis activity of the clarified extract as measured by soluble rhGM-CSF production at small scale as described in Materials and Methods section.

Figure 1a shows that clarified extracts pre-incubated at  $25$ ,  $30$ , or  $37^\circ\text{C}$  show differential amounts of rhGM-CSF are produced in the subsequent 5 h cell-free synthesis reaction. Optimal extract activation was achieved at  $30^\circ\text{C}$  using a 2.5-h pre-incubation period and this was confirmed with extract produced from cells harvested using large scale disc-stack centrifugation, as shown in Figure 1B. All subsequent reactions used extract pre-incubated under these activation conditions.

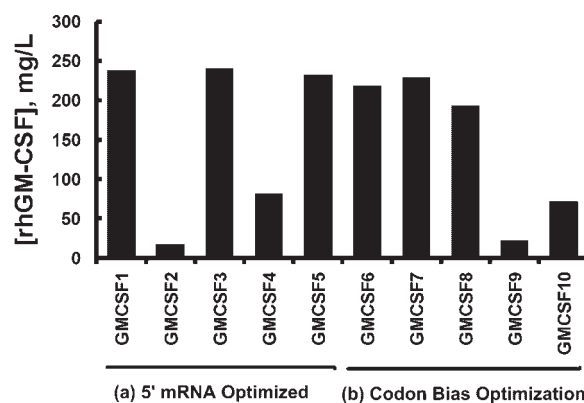
We also investigated the production capacity of activated extract after freezing and storage at  $-40^\circ\text{C}$  and  $-80^\circ\text{C}$  for periods up to 1 year (data not shown). Full rhGM-CSF production activity was retained, even after multiple freeze–thaw cycles, suggesting that active extract can be stored frozen for subsequent rapid protein production.



**Figure 1.** Protein synthesis activity is dependent on the time and temperature of extract pre-incubation. **a:** Aliquots of the cell-free extract were incubated at various times and temperatures (■: 37°C, ●: 30°C, and ▲: 25°C) to optimize the cell-free synthesis activity. **b:** Aliquots of the cell-free extract prepared using a tubular bowl centrifuge (●) or from cells harvested with large scale disc-stack continuous centrifugation, followed by lysate clarification in a tubular bowl centrifuge (○) were incubated at 30°C for various times. The concentration of rhGM-CSF produced after 5 h in a 60  $\mu$ L scale reaction from these various pre-incubation conditions was determined by [<sup>14</sup>C]-leucine incorporation as described in the Materials and Methods section. CVs were  $\leq 10\%$  and the protein samples were  $\geq 95\%$  soluble for all time points.

### Optimization of Gene Sequence Elements

Synonymous gene mutations do not alter the encoded protein, but they can radically influence gene expression (Welch et al., 2009). To investigate how, we designed several gene constructs with synonymous codons designed to alter (a) the instability of the 5' mRNA sequence or (b) codon usage based on the algorithms described by Welch et al. (2009); see Supplementary Methods. As shown in Figure 2, synonymous codon variations lead to variable expression yields. We conclude that either gene sequence optimization method can give reasonable protein yields if multiple gene sequences are screened in the cell-free synthesis reaction. Higher throughput methods based on expression from linear DNA templates may allow a more rapid and extensive screening of this important optimization parameter (Son et al. 2006).

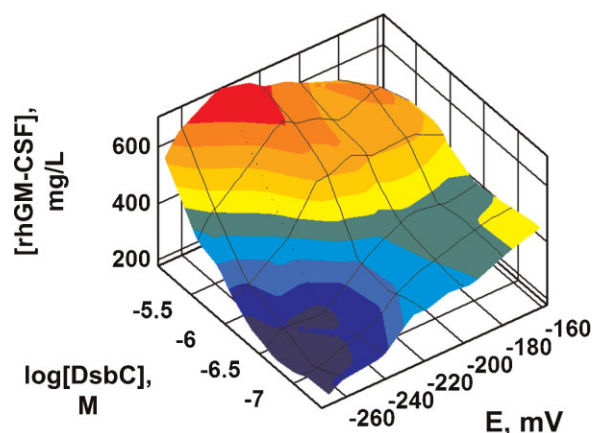


**Figure 2.** Yields of soluble rhGM-CSF for several gene sequences designed using (a) ProteoExpert 5' mRNA optimization or (b) DNA 2.0 codon bias algorithms as described in the Materials and Methods section. The expression of soluble rhGM-CSF was measured by [<sup>14</sup>C]-leucine incorporation in 250  $\mu$ L. CVs were  $\leq 10\%$ , for all constructs tested. The gene sequence of GM-CSF1 (Supplementary Methods) was used to optimize cell-free expression in scale-up experiments.

### Optimization of Disulfide Bond Formation

Conditions of the OCFS reaction can be directly manipulated to provide an optimized environment for both protein expression and folding. In contrast, the ability of *E. coli* to correctly fold recombinant proteins containing multiple disulfide bonds is limited, due the reducing environment of the cytosol and the limited capacity for co-translational folding (Baneyx and Mujacic, 2004; Endo and Sawasaki, 2006; Jermutus et al., 1998; Kawasaki et al., 2003; Kolb et al., 2000; Maier et al., 2005; Yang et al., 2004).

Figure 3 shows that the optimal conditions for production of soluble, folded rhGM-CSF are controlled both by the



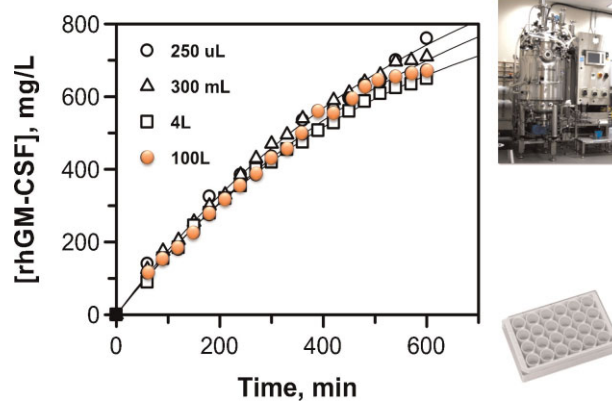
**Figure 3.** Combinatorial optimization of protein folding in the OCFS system. Response surface of redox potential and DsbC concentration on soluble expression of rhGM-CSF. The expression of soluble rhGM-CSF was measured by [<sup>14</sup>C]-leucine incorporation in a 96-well plate for IAM-treated extract as a function of the initial redox potential of added [GSSG]<sub>tot</sub> = 5 mM and DsbC concentration. For reference, the redox potential of the *E. coli* cytosol is  $-270$  mV.

redox potential and the concentration of added *E. coli* disulfide isomerase, DsbC, that catalyzes disulfide bond rearrangement (Yang et al., 2004). These experiments were carried out in 96-well microtiter scale as measured by the amount of soluble rhGM-CSF produced. Under optimized redox conditions the ratio of soluble:total GM-CSF was  $\geq 95\%$ , consistent with correctly folded disulfide bonds, subsequently confirmed by MS analysis of rhGM-CSF produced under these optimized conditions.

Optimization of protein synthesis yield required us to balance efficient translation, which requires reducing conditions (Traut and Haeni, 1967; Yang et al., 2004), with protein folding of correct disulfide bonds that requires a more oxidizing environment. DsbC must be maintained in a reduced state to be able to catalyze the rearrangement of non-native disulfide bonds but has a non-specific folding chaperone activity as well. The high DsbC concentrations required for efficient soluble protein yield suggests that the folding chaperone activity of DsbC dominates, although the drop-off in soluble yield under more oxidizing conditions is consistent with a requirement for reduced DsbC for efficient disulfide isomerase activity (Ryabova et al., 1997).

### OCFS is Linearly Scalable from Microscale to Large Scale Reaction

We used the optimization parameters described above to define process parameters for scale-up, without changes in volumetric performance of the system at all scales. Figure 4 shows that the OCFS process exhibits linear scalability over a  $\sim 10^6$  change in reaction volume for rhGM-CSF, with yields of 700 mg/L after 10 h. This high yield of fully soluble protein makes the requirement for refolding of rhGM-CSF produced in normal bacterial fermentation processes (Berges et al., 1996; Libby et al., 1987) unnecessary



**Figure 4.** The rate of soluble rhGM-CSF production is independent of scale from plates, to stirred tank reactors, to large-scale bioreactors. The initial slopes corresponds to a translation rate of  $\sim 1$  peptide bond per second per ribosome. The concentrations of rhGM-CSF were determined by [ $^{14}\text{C}$ ]-leucine incorporation at 250  $\mu\text{L}$  scale, and by RP HPLC analysis at larger scales, as described in the Materials and Methods section. CVs were  $\leq 10\%$  and the proteins were  $\geq 95\%$  soluble for all time points. Lines are shown for visual purposes only.

and simplifies subsequent downstream processing (vide infra).

### OCFS is Efficient

The overall efficiency of the OCFS system compares favorably with corresponding efficiencies for protein translation in vivo (Table I). Elongation rates of  $\sim 1$  amino acid (aa) per second per ribosome are about 20-fold higher than comparable eukaryotic cell-free systems (Endo and Sawasaki, 2006; Kawasaki et al., 2003), but smaller than the rate of  $10\text{--}20 \text{ aa s}^{-1}$  in vivo for rapidly growing *E. coli*. Similarly, initiation rates in vitro are smaller than estimated

**Table I.** Ribosomal translation in vivo and in vitro<sup>a</sup>

Translation system	Translation efficiency		Productive synthesis time (h)	References
	Initiation rate <sup>b</sup> ( $\text{s}^{-1}$ )	Elongation rate ( $\text{AA s}^{-1} \text{ ribosome}^{-1}$ )		
<i>E. coli</i> cell	0.5	$10\text{--}20^c$	5–24	Gromadski and Rodnina (2004); Kierzek et al. (2001); Parker (1989)
OCFS system	$0.0015^d$	$\sim 1^e$	$\sim 10$	This work
Cytomim system	0.017	$\sim 1$	$\sim 4$	Underwood et al. (2005); Voloshin and Swartz (2008)
PURE system		2	1	Shimizu et al. (2001)
Wheat germ extract (WGE)		$0.05^f$	$\sim 2$	Kawasaki et al. (2003)

<sup>a</sup>At  $37^\circ\text{C}$ , unless otherwise indicated.

<sup>b</sup>Initiation rate is the rate of ribosome clearance, as described by Kierzek et al. (2001).

<sup>c</sup>Translation rate estimated based on in vivo data for *E. coli* in minimal or rich medium or estimated from single-turnover experiments.

<sup>d</sup>Determined from the kinetics of GFP production using the model of Kierzek et al., unpublished results.

<sup>e</sup>Translation rate at  $30^\circ\text{C}$  derived from initial translation rate:  $150 \text{ mg/L/h}$  (Fig. 4), rhGM-CSF MW = 15 kDa, and an active ribosome concentration =  $300 \text{ nM}$  (Supplementary Fig. S5).

<sup>f</sup>Batch mode scFv synthesis using eukaryotic wheat germ extract (WGE) translation. The elongation rate for a 26 kDa scFv is estimated from the observed maximal synthesis rate of  $18 \mu\text{g/mL/h}$  total protein, 85% soluble protein, a WGE  $A_{260} = 42$ ; where  $1 A_{260} = 20 \text{ nM}$  eukaryotic ribosomes. The translation elongation rate is 10 times slower in eukaryotes.

rates in vivo. In vivo expression of heterologous proteins must compete with the protein synthesis required for cell growth and viability, so that the effective heterologous protein synthesis rate in vivo may be lower. The slower translation rate may also aid in efficient co-translational folding of rhGM-CSF in the presence of the directly added DsbC chaperone.

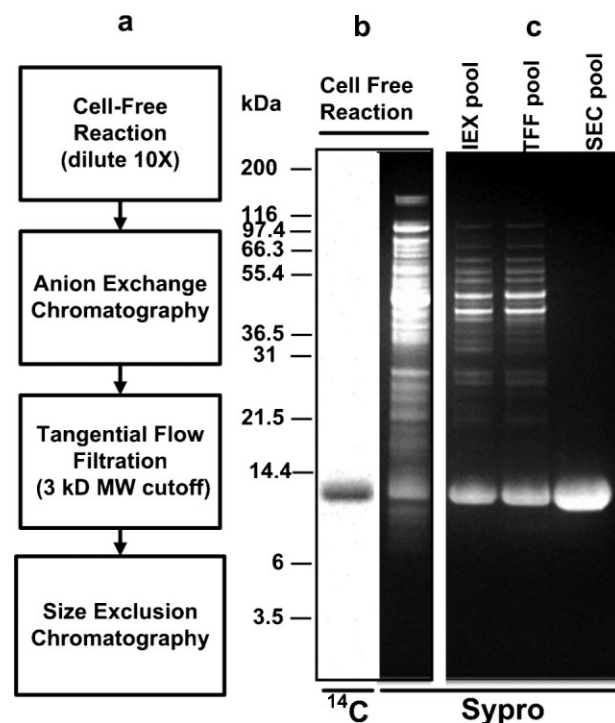
The yields achieved here were obtained using standard stirred tank bioreactors in batch mode without feeding supplemental reagents over the 10 h production process. OCFS is able to produce protein for longer times compared to other in vitro protein synthesis systems in a comparable batch mode (Table I), due to the ability to maintain steady state levels of NTPs generated by active oxidative phosphorylation pathways. As alternatives, continuous or semi-continuous formats of cell-free protein synthesis that require a supply of high-energy NTP substrates and a reservoir to remove inhibitory byproducts have been developed (Endo and Sawasaki, 2006; Jewett and Swartz, 2004b; Spirin et al., 1988). Although continuous-feed approaches increase the duration of protein synthesis, they are not scalable, are expensive, and are unwieldy due to membrane fouling for both high-throughput screening and large-scale production.

### OCFS Allows Efficient Downstream Protein Purification

In engineering an effective large-scale production method for biologic products it is important to design a system that uses standard downstream process equipment and reduces the requirement for multiple manipulations of the product during the purification process. We therefore developed a streamlined two-column process for purification of rhGM-CSF directly from the cell-free production extract, as outlined in Figure 5.

Because the cell-free reaction product pool contains soluble, correctly folded protein product and is substantially reduced in endotoxin levels and other cellular debris, it could be loaded directly onto an anion exchange resin, without significant degradation in column performance during multiple purification runs. Subsequent concentration by tangential flow filtration and removal of host cell proteins by size exclusion chromatography (SEC)-yielded fully bioactive product with 65% overall recovery of high purity rhGM-CSF.

To assess protein quality produced by the OCFS system, we used a series of common comparability protocols in biopharmaceutical manufacturing (Chirino and Mire-Sluis, 2004), as summarized in Figure 6 and Supplementary Methods. We confirmed the absolute mass and correct disulfide bonding patterns by LC-MS, and LC-MS/MS sequencing, SEC was used to determine purity, and the biological activity of rhGM-CSF was determined in a cell-based proliferation assay. The purity achieved could be considered suitable for therapeutic use but could also be improved further using three-step purification. We

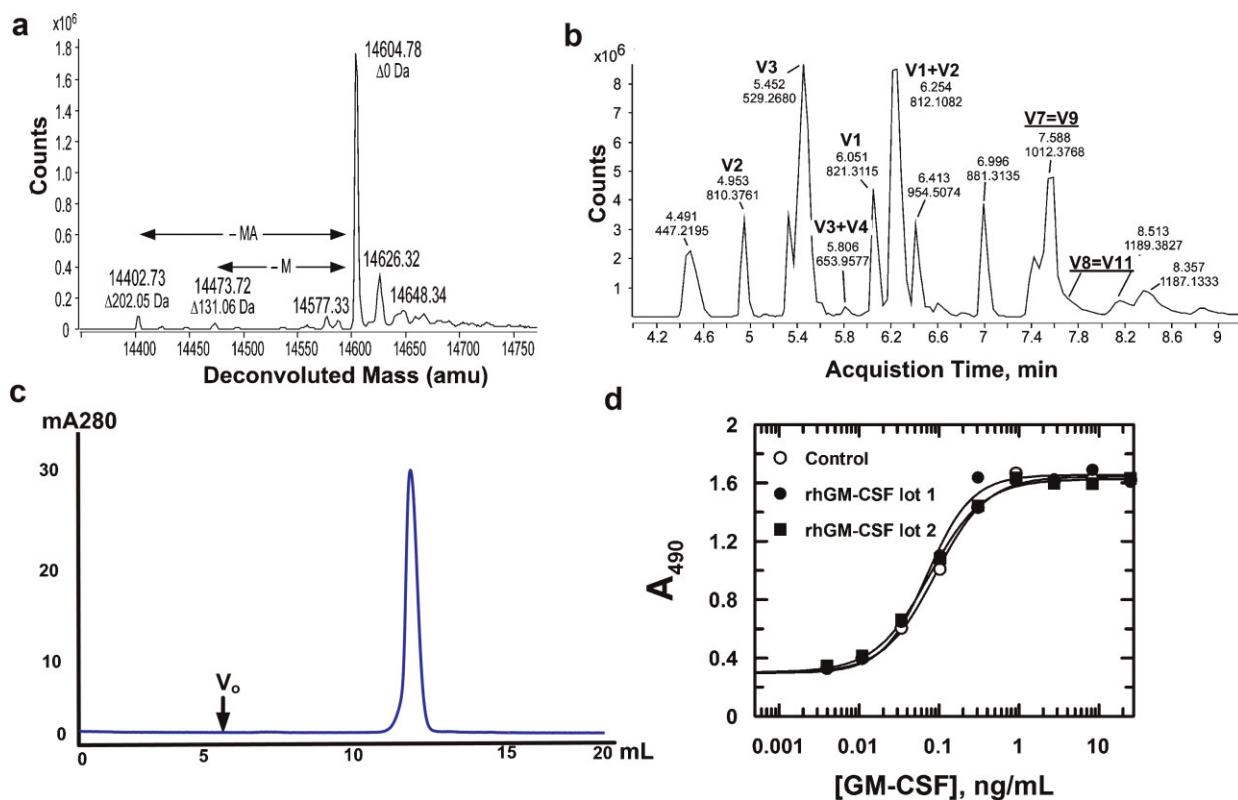


**Figure 5.** Analysis of process purification steps used in the production of pharmaceutical-grade rhGM-CSF. **a:** Process flow diagram for protein purification. **(b)** Cell-free reaction product pool was visualized, using non-reducing SDS-PAGE, by  $^{14}\text{C}$  autoradiography of leucine incorporation at 5 h that measures only protein produced, or by Sypro staining of all proteins in the extract at 10 h. Autoradiography indicates incorporation of  $^{14}\text{C}$ -leucine corresponding to the molecular weight of rhGM-CSF and confirms that only rhGM-CSF was produced because the T7 RNAP limits transcription to the T7 promoter on the plasmid added to the reaction, with no evidence for aberrant protein products produced from native RNAP transcription. **c:** Non-reducing SDS-PAGE of the process purification pools. **Lane 1:** Anion exchange pool. **Lane 2:** Tangential flow filtration pool. **Lane 3:** SEC final product pool.

conclude that OCFS can readily produce proteins of pharmaceutical grade quality at commercially relevant yields and scales.

### Discussion

Several cell-based secretion systems produce biologically active GM-CSF in *E. coli* (Berges et al., 1996; Libby et al., 1987), yeast [50–60 mg/L; (Price et al., 1987)], *Aspergillus niger* [0.64  $\mu\text{g/L}$ ; (Kim et al. 2000)], *Pichia pastoris* (200–250 mg/L; (Sainathan et al., 2005) and 130 mg/L; [Bhattacharya et al., 2007]), and insect cells (11–45 mg/L; [Chiou and Wu, 1990]). In general, substantial losses of product due to refolding, filtration, and extensive column chromatography steps are required to isolate homogeneously pure product in these systems. Cell-free protein synthesis of GM-CSF as well as other small, disulfide-bonded proteins has been described previously at small scale (Goerke and Swartz, 2008; Kawasaki et al., 2003; Ryabova et al., 1997; Son et al., 2006; Yang et al., 2004).



**Figure 6.** rhGM-CSF is correctly folded and functional. **a:** Liquid chromatography–Mass spectrometry analysis of intact mass of 14604.78 Da corresponds to the expected average neutral mass of 14604.75 Da. Two minor product related impurities were observed corresponding to removal of the first two amino acids (-M and -MA). These peak intensities do not correspond accurately to the relative amount of product related impurities. **b:** Base peak chromatogram of a V8 protease digestion of rhGM-CSF. The two peptides with disulfide bridges labeled V7 = 9 and V8 = 11 (underlined) coincide with the theoretical masses with correct disulfide bridge formation. The V7 = 9 linkage was confirmed by tandem MS and the V8 = 11 by accurate mass. Other observed V8 digest peptides are marked accordingly. Digest peptides denoted with a plus sign indicate a missed cleavage. **c:** Analytical size exclusion chromatogram of purified rhGM-CSF on a Tosoh TSKgelG3000SWXL column, flow rate 0.5 mL/min. **d:** The effective dose 50, ED<sub>50</sub> <0.1 ng/ml of rhGM-CSF induced proliferation of human TF-1 cells, corresponds to a specific activity of  $14 \times 10^6$  IU/mg, comparable to commercially available *E. coli*-derived rhGM-CSF protein as a positive control. [Color figure can be seen in the online version of this article, available at <http://wileyonlinelibrary.com/bit>]

Our approach represents an advance towards a robust (>95% correctly folded multi-disulfide-bonded protein), high yield (700 mg/L protein over 10 hr), linearly scalable (~ $10^6$  fold), and high-recovery (>65%) system for industrial production of disulfide-bonded proteins such as GM-CSF. Using the open nature of cell-free synthesis we optimized several factors affecting transcription, translation, and folding to predictably identify conditions for scalable production of a disulfide-bonded protein to the 100 L scale. In this *in vitro* synthetic biology-based system, experimental data built from an understanding of the energy source as well as transcription, translational, and protein folding components are sufficient to guide design and testing of protein expression conditions in a microtiter plate format with predictable properties at any scale. More complex secondary effects associated with *in vivo* gene expression and regulatory networks are avoided. For example, in T7-based systems *in vivo*, transcription, translation, and folding are uncoupled. The uncoupling mediated by T7 RNAP may in turn mediate protein aggregation. However, after full

optimization of plasmid, T7 RNAP, and redox conditions, >95% of the protein produced is soluble, there is no evidence of significant misfolded aggregates formed as measured by non-reducing SDS-PAGE autoradiography (Fig. 5b), or detected in the final product (Fig. 6c).

Previous optimization strategies for cell-free protein expression have sampled parameters sparsely and thus cannot be generalized to the more complex, finely balanced networks of protein synthesis and folding described here. Instead, we found that one must first assemble and experimentally characterize interacting networks of myriad system components at small scale to create generalizable response surfaces such as those in Figures 1 and 3 with quantitative predictive capabilities for process scale-up.

The OCFS system uses a K-12-derived production strain as the basis for a cell-line banked under cGMP conditions. Multiple protein products can now be produced on an as needed basis from the same stored frozen extract produced from this strain. For example, similar scalable results to 4 L have been obtained for a single-chain antibody fragment

(scFv), anti-IL-23 scFv (1 g/L in 10 h) and an anti-IL13R $\alpha$ 1 Fab antibody with five disulfide bonds (300 mg/L in 9 h; manuscript in preparation). Use of a single cell-line coupled with the speed of process optimization inherent in the linearly scalable OCFS protein production system should allow for decreases in the timelines for biopharmaceutical protein production and process development.

We have developed the OCFS production process using standard microbial fermentation and process equipment known to scale to thousands of liters under cGMP standards. Additional efforts to improve this production process toward cGMP compliance include (a) developing methods for sterile filtration of the extract, (b) establishing extract process production reproducibility, (c) establishing the translational fidelity of OCFS produced proteins (Heibeck et al., manuscript in preparation), and (d) accurate cost-modeling of the reagents and other process variables. Finally, we aim to develop a molecular understanding of the limiting factors accounting for the time and temperature dependence of extract activation shown in Figure 1, by quantitative LC-MS-based profiling of proteins and metabolites (Roy and Becker, 2007), in order to improve protein production yields further.

The OCFS system is limited to producing proteins with only a few post-translational modifications such as disulfide bonds and proteolytic processing (Son et al., 2006); other therapeutic protein classes such as glycosylated proteins are currently outside the scope of OCFS at large scale. However, OCFS has the potential for synthesis of proteins containing site-specifically incorporated non-natural amino acids (Noren et al., 1989) without the requirement for orthogonal tRNA synthetase-tRNA pairs or limitations to cellular uptake of non-natural amino acids (Antonczak et al. 2009). This will allow the engineering of novel post-translationally modified proteins using the scalable OCFS system.

Portions of this work were supported by STTR grants NSF0420208 & NSF0522337 from National Science Foundation. We thank Soheyl Tadjiki (Postnova Analytics, Salt Lake City, Utah) for Asymmetrical Flow Field-Flow Fractionation data, and Tom Zupancic (www.Virtual3C.com) for helpful comments and criticisms. The authors are grateful for the efforts of all the people who contributed to the work presented here including Albert Boot, John Garbutt, Eudean Garces, Tyler Heibeck, James Krocak, Seung Lee, Patricio Riquelme, James Rozzelle, Alexei Voloshin, and Juan Zhang.

## References

- Antonczak AK, Simova Z, Tippmann EM. 2009. A critical examination of *Escherichia coli* esterase activity. *J Biol Chem* 284(42):28795–28800.
- Baneyx F, Mujacic M. 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotech* 22(11):1399–1408.
- Berges H, Joseph-Liauzun E, Fayet O. 1996. Combined effects of the signal sequence and the major chaperone proteins on the export of human cytokines in *Escherichia coli*. *Appl Environ Microbiol* 62(1):55–60.
- Bhattacharya P, Pandey G, Mukherjee KJ. 2007. Production and purification of recombinant human granulocyte-macrophage colony stimulating factor (GM-CSF) from high cell density cultures of *Pichia pastoris*. *Bioprocess Biosyst Eng* 30(5):305–312.
- Calhoun KA, Swartz JR. 2006. Total amino acid stabilization during cell-free protein synthesis reactions. *J Biotechnol* 123(2):193–203.
- Chiou CJ, Wu MC. 1990. Expression of human granulocyte-macrophage colony-stimulating factor gene in insect cells by a baculovirus vector. *FEBS Lett* 259(2):249–253.
- Chirino AJ, Mire-Sluis A. 2004. Characterizing biological products and assessing comparability following manufacturing changes. *Nat Biotech* 22(11):1383–1391.
- Endo Y, Sawasaki T. 2006. Cell-free expression systems for eukaryotic protein production. *Curr Opin Biotechnol* 17(4):373–380.
- Goerke AR, Swartz JR. 2008. Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol Bioeng* 99(2):351–367.
- Gromadski KB, Rodnina MV. 2004. Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* 13(2):191–200.
- Jermutus L, Ryabova LA, Pluckthun A. 1998. Recent advances in producing and selecting functional proteins by using cell-free translation. *Curr Opin Biotechnol* 9(5):534–548.
- Jewett MC, Swartz JR. 2004a. Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng* 86(1):19–26.
- Jewett MC, Swartz JR. 2004b. Rapid expression and purification of 100 nmol quantities of active protein using cell-free protein synthesis. *Biotechnol Prog* 20(1):102–109.
- Jewett MC, Voloshin A, Swartz J. 2002. Prokaryotic systems for in vitro expression. In: Weiner MP, Lu Q, editors. *Gene cloning and expression technologies*. Westborough, MA: Eaton Publishing. pp. 391–411.
- Kawasaki T, Gouda MD, Sawasaki T, Takai K, Endo Y. 2003. Efficient synthesis of a disulfide-containing protein through a batch cell-free system from wheat germ. *Eur J Biochem* 270(23):4780–4786.
- Kierzek AM, Zaim J, Zielenkiewicz P. 2001. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J Biol Chem* 276(11):8165–8172.
- Kim MJ, Kwon TH, Jang YS, Yang MS, Kim D. 2000. Expression of murine GM-CSF in recombinant *Aspergillus niger*. *J Microbiol Biotechnol* 10:287–292.
- Knapp KG, Goerke AR, Swartz JR. 2007. Cell-free synthesis of proteins that require disulfide bonds using glucose as an energy source. *Biotechnol Bioeng* 97(4):901–908.
- Kolb VA, Makeyev EV, Spirin AS. 2000. Co-translational folding of an eukaryotic multidomain protein in a prokaryotic translation system. *J Biol Chem* 275(22):16597–16601.
- Libby RT, Braedt G, Kronheim SR, March CJ, Urdal DL, Chiaverotti TA, Tushinski RJ, Mochizuki DY, Hopp TP, Cosman D. 1987. Expression and purification of native human granulocyte-macrophage colony-stimulating factor from an *Escherichia coli* secretion vector. *DNA* 6(3):221–229.
- Liu DV, Zawada JF, Swartz JR. 2005. Streamlining *Escherichia coli* S30 extract preparation for economical cell-free protein synthesis. *Biotechnol Prog* 21(2):460–465.
- Maier T, Ferbitz L, Deuerling E, Ban N. 2005. A cradle for new proteins: Trigger factor at the ribosome. *Curr Opin Struct Biol* 15(2):204–212.
- Michel-Reydellet N, Calhoun K, Swartz J. 2004. Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome. *Metab Eng* 6(3):197–203.
- Noren C, Anthony-Cahill S, Griffith M, Schultz P. 1989. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244(4901):182–188.
- Parker J. 1989. Errors and alternatives in reading the universal genetic code. *Microbiol Mol Biol Rev* 53(3):273–298.
- Price V, Mochizuki D, March CJ, Cosman D, Deeley MC, Klinke R, Clevenger W, Gillis S, Baker P, Urdal D. 1987. Expression, purification and characterization of recombinant murine granulocyte-macrophage colony-stimulating factor and bovine interleukin-2 from yeast. *Gene* 55(2–3):287–293.



- Roy SM, Becker CH. 2007. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling. *Methods Mol Biol* 359:87–105.
- Ryabova LA, Desplancq D, Spirin AS, Pluckthun A. 1997. Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones. *Nat Biotechnol* 15(1):79–84.
- Sainathan SK, Tu L, Bishnupuri KS, Han M, Li A, Newberry RD, McDonald KG, Crimmins DL, Houchen C, Anant S., et al. 2005. PEGylated murine granulocyte-macrophage colony-stimulating factor: Production, purification, and characterization. *Protein Expr Purif* 44(2):94–103.
- Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T. 2001. Cell-free translation reconstituted with purified components. *Nat Biotechnol* 19(8):751–755.
- Son JM, Ahn JH, Hwang MY, Park CG, Choi CY, Kim DM. 2006. Enhancing the efficiency of cell-free protein synthesis through the polymerase-chain-reaction-based addition of a translation enhancer sequence and the in situ removal of the extra amino acid residues. *Anal Biochem* 351(2):187–192.
- Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242(4882):1162–1164.
- Traut RR, Haeeni AL. 1967. The effect of sulfhydryl reagents on ribosome activity. *Eur J Biochem* 2(1):64–73.
- Underwood KA, Swartz JR, Puglisi JD. 2005. Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnol Bioeng* 91(4):425–435.
- Voloshin AM, Swartz JR. 2005. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. *Biotechnol Bioeng* 91(4):516–521.
- Voloshin A, Swartz JR. 2008. Large-scale batch reactions for cell-free protein synthesis. In: Spirin AS, Swartz JR, editors. *Cell-free protein synthesis*. Weinheim: Wiley-VCH. pp. 207–232.
- Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C. 2009. Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS One* 4(9):e7002.
- Wunderlich M, Glockshuber R. 1993. Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci* 2(5):717–726.
- Yang J, Kanter G, Voloshin A, Levy R, Swartz JR. 2004. Expression of active murine granulocyte-macrophage colony-stimulating factor in an *Escherichia coli* cell-free system. *Biotechnol Prog* 20(6):1689–1696.
- Zawada J, Swartz J. 2005. Maintaining rapid growth in moderate-density *Escherichia coli* fermentations. *Biotechnol Bioeng* 89(4):407–415.
- Zawada J, Swartz J. 2006. Effects of growth rate on cell extract performance in cell-free protein synthesis. *Biotechnol Bioeng* 94(4):618–624.
- Zubay G. 1973. In vitro synthesis of protein in microbial systems. *Annu Rev Genet* 7:267–287.