

Cell-Free Protein Production

PRINT

[| CESG Cell-Free Workshop](#) | [CellFree Sciences](#) | [CFSTRC](#) |

The CELL-FREE PROTEIN PRODUCTION Team is headed by John Markley, PhD, and is responsible for high-throughput protein production for structural investigations from a wheat germ cell-free expression system.



Goals for CESG Cell-Free Team

- Produce samples to screen for expression, solubility and, where appropriate, tag cleavage.
- Provide high-throughput, large-scale production of labeled proteins of high analytical quality.
- Process optimization of economic and high-quality protein production.
- Production of project resource materials for the pipeline.

- [Bioinformatics](#)
- [Cell-Free](#)
- [Cloning](#)
- [Crystallography](#)
- [Expression Testing](#)
- [NMR](#)
- [Protein Production](#)
- [Protein Purification](#)
- [Quality Assurance](#)
- [Sesame](#)

Overview of Cell-Free at CESG

One of the major bottlenecks in high-throughput structure determination effort is protein expression. Structural proteomics efforts will require expression and purification of thousands of proteins and/or protein fragments. The ability to obtain labeled proteins is an essential aspect required for rapid progress in structural determinations. Critical examples are:

- The production of U-¹⁵N labeled protein to assess the foldedness and aggregation state.
- The production of U-¹³C;¹⁵N or selectively labeled proteins for NMR structure determinations.
- The production of Se-Met labeled protein for phase determination in X-ray studies.

However, many individual proteins cannot be expressed in soluble form in bacteria and are, therefore, not suitable for *E. coli* cell-based production methodologies. Insolubility arises either from an intrinsic property of a protein (for example, aggregation due to a very hydrophobic patch on the surface) or because the protein is not susceptible to the folding mechanisms in the expression host; in which case there is an aggregation of folding intermediates. These include one-third to one half of prokaryote proteins. This proportion is likely to be higher for eukaryotic proteins, particularly those that comprise multiple domains, those that require cofactors or protein partners for proper folding, or those that require extensive post-translational modification. The development of new systems and strategies capable of synthesizing any desired soluble, labeled protein, or protein fragment on a preparative scale as alternatives *E. coli* cell-based production is one of the most important tasks in biotechnology today.

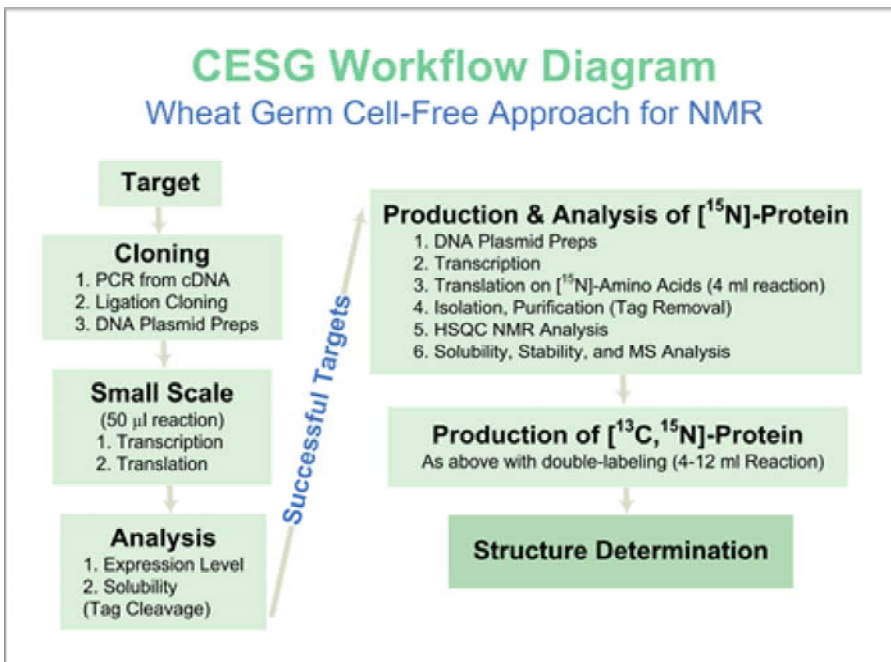
Methods Being Investigated and Publications

+ [Wheat Germ Cell-Free Protein Translation](#)

In cooperation with [Cell-Free Science Technology Research Center at Ehime University \(CFSTRC\)](#) (Matsuyama, Japan) and the [Cell-Free Sciences](#) (Yokohoma, Japan), we are investigating the potential of wheat germ cell-free protein translation as an enabling technology for structural genomics. Two important applications of this technology are a facile assessment of the ability of a given gene to be expressed and the ability to produce isotopically labeled proteins required for certain critical aspects of protein structure determination. A successful implementation of cell-free protein expression may minimize problems in cell harvesting, cell lysis, and pre-column manipulations. Moreover, this approach may simplify

Center for Eukaryotic Structural Genomics - Cell-Free Protein Production

purification efforts as the protein of interest can become the dominant protein in the reaction mixture. Cell-free systems may also permit labeling strategies that cannot be achieved in whole cell systems, while potentially providing a substantial economy in the labeled material required to provide a target protein.



So far, studies carried out at the CESG with this technology on the full-length protein targets from *Arabidopsis thaliana* have yielded the following results:

Screening of 151 targets containing an N-terminal (His)₆-tag on a 50-µL scale gave high-yield expression of highly soluble proteins with a success rate of 50%; screening of 109 targets containing an N-terminal GST-tag resulted in production of highly soluble protein following cleavage of GST fusion tag with PreScission™ Protease (Amersham Biosciences) with a success rate of 49%.

Fifty-two proteins containing an N-terminal (His)₆-tag were expressed on a large scale (4-ml)

using ¹⁵N-labeled amino acids; of these, 46% were obtained in quantities (0.6 mg/ml average yield of purified protein) sufficient for ¹H-¹⁵N HSQC NMR screening; with 58% of those proteins were folded, non-aggregating. Eight folded proteins have been expressed in mg quantities using ¹³C; ¹⁵N-labeled amino acids; from these, two three-dimensional structures have been solved, four others are in progress, and two proteins were obtained in yields insufficient for structure determination. Six proteins containing an N-terminal GST-tag were expressed on large scale (4-ml) using ¹⁵N-labeled amino acids; of these, 100% were obtained in quantities (0.5 mg/ml average yield of purified target protein) sufficient for ¹H-¹⁵N HSQC NMR screening; 33% of those proteins were folded, non-aggregating. One of these folded proteins has been expressed in mg quantities using ¹³C; ¹⁵N-labeled amino acids; one structure is in progress.

+ Advantages and Disadvantages of Wheat Germ Cell-Free Approach

Advantages

- Wheat germ cell-free method supports rapid and efficient screening (supported by robotics).
- Wheat germ cell-free method requires smaller volumes. (avoids lengthy concentration steps in protein purification).
- Labeled proteins can be prepared rapidly (in 1-2 days) to meet needs of structural biologists.
- Supports labeling strategies that are not practical for proteins produced from bacterial cells (no label scrambling).
- Supports the production of eukaryotic N-terminal (His)₆ proteins (previous experience showed that these were not produced successfully from *E. coli* cells).

Disadvantages

- Reagent intensive.
- Currently not compatible with Gateway cloning technology used in other parts of the CESG project.

+ Wheat Germ Cell-Free Instrumentation

CESG will be using the following instrumentation:

Center for Eukaryotic Structural Genomics - Cell-Free Protein Production

CESG utilizes wheat germ cell-free instrumentation from [CellFree Sciences](#) (Yokohama, Japan). Pictured at right are the **GeneDecoder1000 Fully Automated Protein Synthesizer** (LEFT) and the **Protemist Protein Synthesizer** (RIGHT). Both systems make use of wheat germ translation technology developed by CESG collaborator Professor Yaeta Endo, [Director of the Cell-Free Science Technology Research Center at Ehime University](#) (Matsuyama, Japan). The GeneDecoder1000, which performs steps needed for transcription (DNA to mRNA) and translation (mRNA to protein) in 96-well format, enables up to 384 50-microliter reactions to be



carried out overnight. CESG uses this instrument to screen genes or gene fragments to determine the level of protein production, whether a fusion product is cleavable, and the fraction of protein product that is soluble. The Protemist Protein Synthesizer automates up to eight 5-ml translation reactions used in producing proteins for structural investigations. In some cases, a single 5-ml reaction produces sufficient protein for an NMR structure determination. Proteins that express at lower levels require up to three 5-ml reactions.

+ [Head-to-Head Comparison of Protein Production from *E. coli* Cells and Wheat Germ Cell-Free Extracts](#)

CESG has carried out a detailed comparison of its protein production pipelines based on *E. coli* cells and on wheat germ cell-free translation. The *E. coli* cells portion of the experiment made use of CESG's standard maltose binding fusion containing a (His)₆-tag and a TEV protease cleavage site. All targets found to be produced as a soluble, cleaved product were prepared with uniform [¹⁵N]-labeling. The cell-free part of the experiment compared two constructs for each of the 96 targets: one with a non-cleavable N-terminal (His)₆-tag, and one with a cleavable (PreScission Protease, Amersham Biosciences) N-terminal GST-tag. These constructs were first screened on a small scale (50-mL) to determine the level of protein produced and its solubility. Targets that produced soluble protein were then produced as [U-¹⁵N]-proteins in larger scale (4 to 12 mL) cell-free translation reaction mixtures that contained ¹⁵N-labeled amino acids. The proteins produced by each method were analyzed by ¹H-¹⁵N correlation NMR spectroscopy for folding, aggregation state, and stability.

This project, which was completed in August, 2004, has provided a rich source of information. We are only beginning to mine the results to learn what they tell us about these different approaches. When the success rates of individual steps, supplies, and labor are taken into account, the costs for making labeled proteins for NMR structure determinations by the two platforms (*E. coli* cells and wheat germ cell-free) are equivalent. The potential advantage of the cell-free approach is that nearly twice as many of the protein targets in this study prepared by cell-free than by *E. coli* cells yielded samples suitable for NMR structure determination. However, when the *E. coli* cell-free pipeline works, the yields of labeled proteins are higher.

+ [Wheat Germ Cell-Free Protein Production and Stable-Isotope Labeling Platform for NMR-Based Structural Proteomics](#)

In collaboration with Professor Yaeta Endo (Ehime University, Matsuyama, Japan) and [CellFree Sciences](#) (Yokohama, Japan), CESG has developed a platform that utilizes wheat germ cell-free technology to produce protein samples for NMR structure determinations. In the first stage, cloned DNA molecules coding for proteins of interest are transcribed and translated on a small scale (25 microL) to determine levels of protein expression and solubility. The amount of protein produced (typically 2-10 microgram) is sufficient to be visualized by polyacrylamide gel electrophoresis. The fraction of soluble protein is estimated by comparing gel scans of total protein and soluble protein. Targets that pass this first screen by exhibiting high protein production and solubility move to the second stage. In the second stage, the DNA is transcribed on a larger scale, and labeled proteins are produced by incorporation of [¹⁵N]-labeled amino acids in a 4 mL translation reaction that typically produces 1-3 mg of protein. The [¹⁵N]-labeled proteins are screened by ¹H-¹⁵N HSQC NMR spectroscopy to determine whether the protein is a good candidate for solution structure determination. Targets that pass this second screen are then translated in a medium containing amino acids doubly labeled with ¹⁵N and ¹³C. These steps

Center for Eukaryotic Structural Genomics - Cell-Free Protein Production

can be automated so that the labor costs involved are minimal. CESC uses an automated platform for wheat germ cell-free production of labeled proteins. Our current robotic systems from [CellFree Sciences](#) (Yokohoma, Japan) include the GeneDecoder1000 (2-5 ug per well in 96-well format), the Protelist10 and the Protelist100 (1-2 mg per sample in eight samples format), and Protelist DT-II (0.1-0.3 mg purified protein per well in 6-well format).

The GeneDecoder1000 is used to produce samples to screen for expression, solubility and, where appropriate, tag cleavage. The Protelist10 and the Protelist100, coupled with ACTA PRIME purification systems, are used for expression and purification of sufficient quantities of labeled protein for NMR structural studies. Our cumulative experience with cell-free expression includes over 1000 different structural genomics targets from human, mouse, *Plasmodium*, and *Arabidopsis*. To date, CESC has deposited into the PDB 23 NMR structures of eukaryotic proteins produced by wheat germ cell-free methodology. The average yield of labeled purified proteins has been ~1.2 mg per ml of wheat germ lysate (OD260=200). We also report that the Protelist DT-II provides a cost-effective and rapid method for screening multiple constructs engineered to improve solubility or foldedness.

Furthermore, CESC has begun structural investigations of membrane proteins using the automated translation and purification capabilities of the Protelist DT-II. Several detergents have been identified to be compatible with wheat germ cell-free translation, and current efforts are aimed at developing efficient ways of protein concentration, detergent exchange, and preparation of structural samples (unpublished results).

[+ CESC Publications on Wheat Germ Cell-Free Methodologies and Technologies](#)

Vinarov, D.A., Loushin Newman, C.L., Tyler, E.M., Markley, J.L. (2006) **Wheat germ cell-free expression system for protein production.** *Current Protocols in Protein Science*, Wiley Interscience.

Vinarov, D.A., Lytle, B.L., Peterson, F.C., Tyler, E.M., Volkman, B.F., Markley, J.L. (2004) **Cell-free protein production and labeling protocol for NMR-based structural proteomics.** *Nat Methods* 1(2):149-53. |[15782178](#)|

Vinarov, D.A., Markley, J.L. (2005) **High-throughput automated platform for nuclear magnetic resonance-based structural proteomics.** *Expert Rev Proteomics* 2(1):49-55. |[15966852](#)|

Vinarov, D.A., Newman, C.L., Markley, J.L. (2006) **Wheat germ cell-free platform for eukaryotic protein production.** *FEBS J* 273(18):4160-9. |[16930128](#)|

[+ Production of Se-Met Labeled Proteins for X-ray Crystallography by Wheat Germ Cell-Free](#)

We have shown that the wheat germ cell-free protein production platform used by CESC can be used for efficient production of [Se-Met]-labeled proteins with incorporation efficiency close to 100%. The procedure from DNA to purified protein can be carried out in automated fashion on the [CellFree Sciences](#) Protelist DT-II bench-top robot in sufficient quantities for small-scale crystallization screening.

Cell-Free Team

- John Markley, PhD (PI)
- Shin-ichi Makino, PhD

Center for Eukaryotic Structural Genomics (CESG), Department of Biochemistry
University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706 | [Map](#) |
Telephone: 608.263.2183 Fax: 608.890.1942 Email: cesginfo@biochem.wisc.edu

| [Biochemistry](#) | [Private](#) | [PDB](#) | [PSI](#) | [PSI-KB](#) | [PSI-MR](#) | [Sesame](#) | [UW-Madison](#) |
Copyright © 2006 The Board of Regents of the University of Wisconsin System