Center for Eukaryotic Structural Genomics

Protein Structure Initiative

Protein Production for Structural Investigations Based on the Wheat Germ Cell-Free Expression System

Dmitriy A. Vinarov, Ejan M. Tyler, Mark N. Shahan, and John L. Markley

University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin, USA 53706-1549, http://www.uwstructuralgenomics.org

Abstract

In cooperation with Ehime University and CFS, Ltd., we are investigating the potential of an automated high-throughput protein production platform, based on the wheat germ cell-free expression system, as an enabling technology for structural genomics. Studies carried out at the CESG with this technology on full-length protein targets from Arabidopsis thaliana have yielded the following results. (1) Screening of 143 targets containing an N-terminal (His)6-tag on a 50-μL scale gave high-yield expression of proteins with a success rate of 76% and with 65% of expressed proteins being soluble; screening of 106 targets containing an N-terminal GST-tag resulted in high-yield expression with a success rate of 78% and with 65% of the expressed fusion proteins being soluble and 98% of those target proteins being soluble after cleavage of the GST-tag. (2) 48 proteins with uncleavable N-terminal (His)6-tag were expressed in mg quantities with incorporation of ^{15}N-labeled amino acids for ^{1}H-^{15}N HSQC NMR screening; of these, 14 proteins were found to be folded, 8 unfolded, and 26 were produced at yields insufficient for NMR analysis. (3) 6 of these folded proteins have been expressed in mg quantities with incorporation of ^{13}C;^{15}N-labeled amino acids; from these, 2 three-dimensional structures have been solved, and 4 others are in progress. (4) Ni-HiTrap and gel-filtration columns (where appropriate) have been used successfully to purify proteins with N-terminal (His)6-tags.

Rationale for Exploring This Technology

Structural proteomics projects require the expression and purification of thousands of proteins and/or protein fragments. The ability to obtain labeled proteins is an essential requirement for rapid progress in structure determinations. Critical examples are (1) the production of U-^{15}N labeled protein to assess the foldedness and aggregation state, (2) the production of U-^{13}C;^{15}N or selectively labeled proteins for NMR structure determinations, and (3) the production of SeMet labeled protein for phase determination in X-ray studies. The successful implementation of cell-free protein expression may minimize problems in cell harvesting, cell lysis, and pre-column manipulations. Moreover, this approach may simplify purification, because the protein of interest is isolated from a smaller set of contaminants. 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 produce a target protein. This promise warrants serious consideration of how cell-free translation can contribute to high-throughput structural biology.

Potential Advantages of Cell-Free Protein Synthesis Over Other Approaches

- Expression trials can be carried out without sub-cloning
- Improved yields of folded, soluble protein
- Amenable to automation since only small volumes are required
- Simplified protein purification (fewer contaminants, proteases)
- Efficient incorporation of labeled amino acids without label scrambling
- Expression of proteins toxic to cells or degraded in cells
- Post-translation modifications and S-S bond formation may be possible

CESG's Wheat Germ Cell-Free Protein Expression Project Involves Three Major Groups

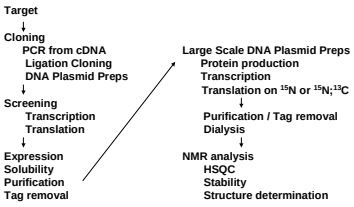
University of Wisconsin-Madison

- Screening of potential targets from eukaryotic genomes for suitability for structural studies
- Production of labeled proteins on the scale of several milligrams
- Assessment of this approach for high-throughput structure determination
- Improvement of the technology through its use in a production environment

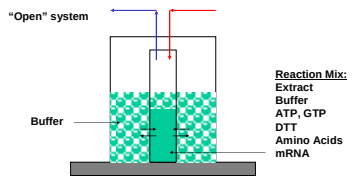
Ehime University and Cell-Free Sciences Co. Ltd.

- Wheat-germ extract production
- Enabling methodology
- Robotics

Work-Flow Diagram of Wheat Germ Cell-Free Protein Expression



Semi-Continuous Wheat Germ Cell-Free Translation System



Buffer contains an energy-regenerating system (phosphocreatine / phosphocreatine kinase), spermidine (to stimulate the efficiency of chain elongation), magnesium acetate, potassium acetate and amino acids.

Expression and Solubility Screening of *Arabidopsis* ORFs with N-terminal $(His)_6$ Tags

EXPRESSION			SOLUBILITY		
Yes	116	77%	>50%	75	65%
No	35	23%	<50%	41	35%
Total	151	100%	Total	116	100%

Overall success rate in expressing soluble protein with N-terminal $(His)_6$ tag was 50%

Expression and Solubility Screening of *Arabidopsis* ORFs with N-terminal GST Tags

EXPRESSION			SOLUBILITY		
Yes	86	79%	>50%	56	65%
No	23	21%	<50%	30	35%
Total	109	100%	Total	86	100%

55 / 56 target proteins remained > 98% soluble after tag cleavage

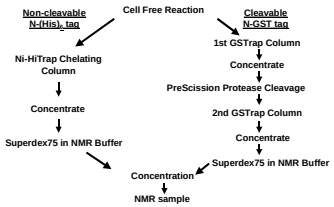
Overall success rate in expressing soluble protein by the cleavable GST route was 49 %

Structural Investigations

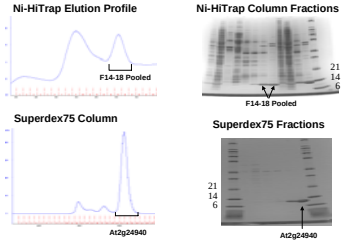
	N-$(His)_6$ Tag		N-GST Tag	
Folded	14	29%	1	33%
Unfolded	9	17%	2	67%
Low Yield	26	54%		
Number of Proteins Expressed	49	100%	3	100%
Average Yield*	0.6 mg/ml		0.75 mg/ml Target protein	
Structures Solved	2			
Structures in progress	4			

*For proteins used in structural investigations

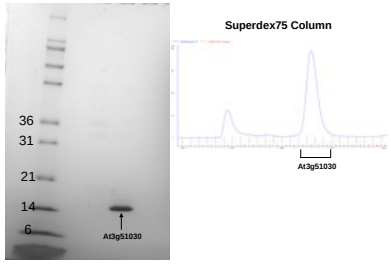
Work-Flow Diagram of Wheat Germ Cell-Free Protein Purification



His-Tag Purification of At2g24940 (12 kDa)



His-Tag Purification of At3g51030 (14 kDa)



Wheat Germ Cell-Free Structure Gallery

At3g01050
12 kDa

At2g24940
14 kDa

Fully Automated Protein Synthesizer

Two Modes of Operation for the GeneDecoder 1000

Screening

- 4 x 96-Well Plates
- Overnight Run
- 2-10 μg Protein / Well
- Uses 2.5 – 5 mL of Wheat Germ Extract / Plate

Small-Scale Protein Production

- 2 x 96-Well Plates
- Overnight Run
- 10-50 μg Protein / Well
- Uses 5 – 10 mL of Wheat Germ Extract / Plate

People Involved

University of Wisconsin-Madison

Cell-Free Protein Production: John Markley, Dmitriy Vinarov, Ejan Tyler, Mark Shahan; NMR: Min Lee, Claudia Cornilescu, Shanteri Singh, Rob Tyler, Jikui Song

Medical College of Wisconsin, Milwaukee/NMR

Brian Volkman, Betsy Lytle, Francis Peterson

Ehime University and Cell-Free Sciences Co. Ltd., Japan

Yaeta Endo, Tatsuya Sawasaki, Motoo Watanabe

CESG is supported by the National Institute of General Medical Sciences through the Protein Structure Initiative (P50 GM 64598).