A wheat germ cell-free system is a novel way to screen protein folding and function

EUGENE HAYATO MORITA,^{1,2} TATSUYA SAWASAKI,^{2,3} RIKOU TANAKA,⁴
YAETA ENDO,^{2,3} AND TOSHIYUKI KOHNO⁴

¹Center for Gene Research, Ehime University, Ehime 790-8566, Japan

²Satellite Venture Business Laboratory, and ³Department of Applied Chemistry, Faculty of Engineering, Ehime University, Ehime 790-8577, Japan

⁴Mitsubishi Kagaku Institute of Life Sciences (MITILS), Tokyo 194-8511, Japan

(RECEIVED November 27, 2002; FINAL REVISION March 4, 2003; ACCEPTED March 4, 2003)

Abstract

For high-throughput protein structural analysis, it is indispensable to develop a reliable protein overexpression system. Although many protein overexpression systems, such as that involving *Escherichia coli* cells, have been developed, the number of overexpressed proteins showing the same biological activities as those of the native proteins is limited. A novel wheat germ cell-free protein synthesis system was developed recently, and most of the proteins functioning in solution were synthesized as soluble forms. This suggests the applicability of this protein synthesis method to determination of the solution structures of functional proteins. To examine this possibility, we have synthesized two ¹⁵N-labeled proteins and obtained ¹H-¹⁵N HSQC spectra for them. The structural analysis of these proteins has already progressed with an *E. coli* overexpression system, and ¹H-¹⁵N HSQC spectra for biologically active proteins have already been obtained. Comparing the spectra, we have shown that proteins synthesized with a wheat germ cell-free system have the proper protein folding and enough biological activity. This is the first experimental evidence of the applicability of the wheat germ cell-free protein synthesis system to high-throughput protein structural analysis.

Keywords: Wheat germ; cell-free; protein synthesis; HSQC; structural analysis

With the increase in the available sequence information on the genomes in various cells, attention has been turned to the structures, properties, and functional activities of proteins. However, rapid progress in the area of proteomics requires the availability of sufficient amounts of proteins. Currently, three major strategies are being used for protein production: chemical synthesis, in vivo expression, and cell-free synthesis. The first two methods have severe draw-

backs. Chemical synthesis is not practical for the synthesis of long peptides (Blaschke et al. 2000), and in vivo expression can produce proteins that do not have any significant effect on the physiology of the host cells (Golf and Goldberg 1987; Chrnyk et al. 1993). With a cell-free translation system, in contrast, one can synthesize larger proteins at the same or higher speed, and as accurately as ones for in vivo translation (Kurland 1982; Pavlov and Ehrenberg 1996), and express proteins that would interfere with the host cell physiology.

One of the most convenient eukaryotic cell-free translation systems is based on wheat germ embryos containing all the components for translation in a concentrated dried state and ready for protein synthesis after germination. A past study has indicated that such systems are generally unstable and thus insufficient (Roberts and Paterson 1973). Recently, however, we found that plants contain endogeneous inhibi-

Reprint requests to: Eugene Hayato Morita, Center for Gene Research, Ehime University, 3-5-7 Tarumi, Ehime 790-8566, Japan; e-mail: ehmorita@dpc.ehime-u.ac.jp; fax: 81-89-946-9968; or Toshiyuki Kohno, Mitsubishi Kagaku Institute of Life Sciences (MITILS), 11 Minamiooyama, Machida-shi, Tokyo 194-8511, Japan; e-mail: tkohno@libra.ls.m-kagaku.co.jp; fax: 81-42-724-6296.

Abbreviation: HSQC, heteronuclear single quantum correlation spectroscopy.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.0241203>.

tors of translation, and that in the case of conventional wheat germ extracts, the RNA N-glycosidase tritin and other inhibitors, such as thionin, ribonucleases, deoxyribonucleases, and proteases—those are found in the endosperm—inhibit translation (Ogasawara et al. 1999; Madin et al. 2000). Extensive washing of wheat embryos, to eliminate endosperm contamination, results in extracts with high degrees of stability and activity (Madin et al. 2000). With such an extract, the translation reaction proceeds for longer than 60 h. When performed in a dialysis bag with continuous feeding of substrates and removal of small byproducts (Spirin et al. 1988), enzymatically active proteins are yielded in milligram quantities per milliliter reaction volume (Madin et al. 2000). In the previous study, Sawasaki et al. (2002) showed the applicability of this cell-free protein synthesis system to the screening of gene products, and suggested that the expressed proteins have functions and then may attain the correct 3D structures. This applicability is crucial for modern proteomics that require a high throughput. However, to clarify the interrelationships between the structures and functions of proteins, it is indispensable to confirm that the functionally active proteins synthesized with this cell-free system have the proper protein folding. In this article, we report that biologically active proteins have the same solution structures as already determined. We also show that with the use of isotopically labeled amino acids as substrates, only the proteins synthesized with this cell-free system are isotopically labeled. This indicates further applicability of this cell-free system because the structural analysis of proteins can progress without any purification step when the concomitant proteins do not exhibit any strong interaction with the synthesized proteins. This property will greatly facilitate structural studies on proteins and is strongly related with the high throughput necessary for modern proteomics. The above-mentioned properties constitute the first experimental evidence that a wheat germ cell-free protein synthesis system is the best way to clarify the structure–function interrelationships of proteins.

Results

Figure 1 shows the amounts of ubiquitin and RbpA1 synthesized in 1- μ L reaction mixtures determined by SDS-PAGE. The synthesized proteins are indicated by arrows. The amounts of the synthesized proteins increased significantly until 4 and 2 days incubation for ubiquitin and RbpA1, respectively. On the basis of these results, we synthesized ^{15}N labeled ubiquitin and RbpA1 for 4 and 2 days with the cell-free system, respectively. The total amounts of the synthesized proteins in 1-mL reaction mixtures were from 200 to 400 μ g (from 20 to 40 nmole) and the final concentrations of NMR samples were ~ 40 μM . In Figures 2 and 3, the ^1H - ^{15}N HSQC spectra of proteins overexpressed in *E. coli* cells and those synthesized with the wheat germ

Figure 1. SDS-PAGE of reaction mixtures for (A) ubiquitin and (B) RbpA1, after synthesis. The synthesized proteins are indicated by black (for ubiquitin) and white (for RbpA1) arrows.

cell-free system are compared. On the synthesis of ubiquitin with the cell-free system, the N atoms in the side chains of Asn and Gln were not ^{15}N labeled, and the corresponding ^{15}NH signals were not observed. Almost all the backbone ^{15}NH signals overlapped. In the case of RbpA1 synthesis, on the other hand, the N atoms in the side chains of Asn and Gln were ^{15}N labeled and almost all the ^{15}NH signals overlapped. In both cases, it is indicated that the overall structures of proteins synthesized in the two different ways are almost identical.

Furthermore, the ^1H - ^{15}N HSQC spectra of proteins overexpressed in *E. coli* cells were for purified proteins, and those of proteins synthesized with the cell-free system for crude ones. This difference in sample conditions indicates that the concomitant proteins in the reaction mixture are not ^{15}N labeled in the process of targeted protein synthesis.

Discussion

In Figure 1, bands corresponding to the synthesized proteins can be clearly observed, and it can be estimated that the amounts of the synthesized proteins are from 200 to 400 ng/ μL on the basis of the intensities of these bands. Then, the total amount of the synthesized proteins in 1-mL reaction mixture can be estimated to from 200 to 400 μg . The molecular weights of the synthesized proteins are almost 10,000, and the molar amounts of the synthesized proteins are from 20 to 40 nmole. These results are almost the same as the previous ones (Sawasaki et al. 2002).

Next, as shown in Figure 2, the ^1H - ^{15}N HSQC spectra for the overexpressed and purified proteins are almost the same

Figure 2. ^1H - ^{15}N HSQC spectra (NMR buffer, pH 6.0, 30°C) of ^{15}N -labeled yeast ubiquitin overexpressed in *E. coli* cells (1.0 mM, 128 [t1] $\times$ 1024 [t2] complex points, 64 scans; black) and synthesized with the wheat germ cell-free system (0.10 mM, 64 [t1] $\times$ 512 [t2] complex points, 512 scans; red) were obtained at the ^1H resonance frequency of 500 MHz. To optimize the resolution in the nitrogen dimension, a ^{15}N spectral width of 1600 Hz was used (spectral widths of 1600 and 6250 Hz in F_1 and F_2 , respectively).

as those for ones synthesized with the wheat germ cell-free system and crude proteins. This reveals an important feature of the proteins synthesized with the wheat germ cell-free system. In the ^1H - ^{15}N HSQC spectra for the reaction mixtures, only signals corresponding to the newly synthesized proteins can be observed. This means that the amino acids added to the dialysis buffer as substrates are used to synthesize the target proteins following the genetic information of mRNA added to the reaction mixture. In the case of the cell-free protein synthesis system involving *E. coli* extracts, without purification, some signals of concomitant impurities or structural heterogeneity of synthesized proteins will be observed in the ^1H - ^{15}N HSQC spectra of reaction mixtures (data not shown). Shimizu et al. (2001) recently reconstituted a cell-free translation system with purified components from *E. coli* extracts. However, the applicability of

this system to protein structural analysis has not been shown yet. The presence of these contaminating signals may interfere with checking of the folding of synthesized proteins or monitoring of the molecular interactions between the synthesized proteins and substrates. In our case, without purification, no contaminating signal was observed in ^1H - ^{15}N HSQC spectrum, and this feature of the wheat germ cell-free system will greatly facilitate the screening of synthesized-protein folding and determination of protein structures, compared to the case in which proteins are synthesized or overexpressed with other systems. This is one of the important features for high-throughput proteomics.

In Figure 3, the ^1H - ^{15}N HSQC spectra of purified RbpA1 synthesized in two different ways can be seen to be quite identical. Without purification, the ^1H - ^{15}N HSQC spectrum of RbpA1 synthesized with the wheat germ cell-free system

Figure 3. ^1H - ^{15}N HSQC spectra (NMR buffer, pH 6.9, 30°C) of ^{15}N -labeled RbpA1 overexpressed in *E. coli* cells (0.5 mM, 128 [t1] $\times$ 512 [t2] complex points, 64 scans; black) and synthesized with the wheat germ cell-free system (0.12 mM, 64 [t1] $\times$ 512 [t2] complex points, 1024 scans; red) were obtained at the ^1H resonance frequency of 500 MHz. To optimize the resolution in the nitrogen dimension, a ^{15}N spectral width of 1500 Hz was used (spectral widths of 1500 and 8000 Hz in F_1 and F_2 , respectively).

is much weaker and different from these spectra. This situation is drastically changed by the treatment of a crude RbpA1 sample with RNaseA. After such treatment, the two spectra were almost identical, which indicates that the RbpA1 interacts with the nucleic acid molecules present in the reaction mixture for the wheat germ cell-free system. However, in the case of ubiquitin, this tendency was not observed. These results revealed other important features. In the absence of concomitant molecules showing great interaction with the target proteins, the ^1H - ^{15}N HSQC spectra of proteins overexpressed in *E. coli* and those of synthesized proteins are identical. For the proteins examined in this study, it was shown that the biological activities of these overexpressed in *E. coli* are as high as those of the native proteins. Considering these results, it can be concluded that the proteins synthesized with the wheat germ cell-free system have the proper protein folding and enough biological activity. It is further concluded that if the ^1H - ^{15}N HSQC spectral pattern for a synthesized functionally unidentified

protein changes during the purification steps, this protein must strongly interact with the concomitant molecules, which can be extracted in these steps. This means that it is easy to determine the biological function of this protein and the target molecules. From this point of view, our results can be interpreted as follows.

Ubiquitin does not interact with the concomitant molecules in a reaction mixture, and no significant spectral changes are observed between the ^1H - ^{15}N spectra for purified and crude solutions. However, in the case of RbpA1, RbpA1 interacts with RNA molecules electrostatically, and the signal intensities observed in the ^1H - ^{15}N HSQC spectrum are quite weak because of the increase in the apparent molecular weight. After ultracentrifugation and polyethylene imine treatment, the level of concomitant RNA molecules becomes lower and the signal intensities observed in the ^1H - ^{15}N HSQC spectrum increase significantly. From this result, RbpA1 is thought to interact with RNA molecules strongly. As shown in a previous study (Sato 1995),

RbpA1 has inhibits the formation of double-stranded mRNA, and the spectral change observed here is in good agreement with this biological function of RbpA1.

Next, we will focus our attention to the cost of this protein synthesis. We purchased the wheat germ extract from Cell-Free Science Company. Concentration of the amino acids needed to synthesize the protein with wheat germ cell-free system is about half of that with *E. coli* cell-free system (RTS500; Roche). The total amounts obtained in 1-mL reaction mixtures with these systems are almost the same. Then, the total cost needed to synthesize fully ^{15}N -labeled proteins with wheat germ cell-free system is almost the same as that with RTS500 system. To synthesize the ^{15}N and ^{13}C double-labeled proteins, the price of each double-labeled amino acid is quite expensive, and the total cost of the *E. coli* system will be much higher than that of the wheat germ system.

In the wheat germ system, the ^{15}N -labeled amino acid mixture from the algal cell can be used as the amino acid source. In this mixture, some types of amino acids are not included, and we have to add these amino acids before use. However, the price of this mixture is quite cheap (less than a 10th of the total cost of pure ^{15}N -labeled or ^{15}N - and ^{13}C -double-labeled amino acids). Then, in the case of a uniform labeling experiment (^{15}N , ^{15}N and ^{13}C), economically, we must use the amino acid mixture from algal cells as the labeled amino acid source.

In summary, our results strongly indicate that the wheat germ cell-free protein synthesis system is one of the best ways for both protein structural analysis and the screening of the biological functions of newly found proteins.

One other important feature of the cell-free system is selective amino acid labeling of proteins. We now try to check this possibility, and despite the incompleteness of our checking, amino acids checked until now have been selectively labeled in proteins with our wheat germ system. In the near future, we will show precise information in another article.

Materials and methods

Synthesis of mRNA

The coding sequences of yeast ubiquitin and RbpA1 were amplified by the PCR method, and transferred to the EcoRV and XhoI sites of the pEU3-NII plasmid (Toyobo; T7 promoter sequence is exchanged with SP6 promoter sequence). In the presence of 16 mM Mg^{2+} , the mRNAs of these proteins were synthesized with SP6 RNA polymerase, with these plasmids as templates (Madin et al. 2000; Sawasaki et al. 2002).

Protein synthesis

Synthesized mRNAs (200 μg) were precipitated with ethanol and dissolved in 260 μL dialysis buffer, and then mixed with the wheat

germ extract for protein synthesis (Madin et al. 2000). This mixture was dialyzed against the dialysis buffer containing 20 amino acids labeled with ^{15}N (Nippon Sanso) for 4 days (ubiquitin) or 2 days (RbpA1). Wheat germ extract was purchased from the Cell-Free Science Company, and other reagents (other than labeled amino acids) were purchased from Nakarai Tesque. After synthesis, the reaction mixtures were treated in two different ways, as follows.

Ubiquitin

The reaction mixture (1 mL) was concentrated to 250 μL with the use of a Centricon-3 micro-concentrator (Millipore), then the buffer was changed to NMR buffer (50 mM sodium phosphate, 100 mM NaCl, pH 6.5) by passage through a Micro Spin G-25 column (Pharmacia) equilibrated with the same buffer.

RbpA1

The reaction mixture (6 mL) was subjected to ultracentrifugation (100,000g for 1 h), and then polyethylene imine was added to the supernatant to the final concentration of 0.5% to remove the concomitant nucleotide derivatives. This solution was subjected to centrifugation (20,000g for 15 min), and then ammonium sulfate was added to the supernatant to the final concentration of 430 mg/mL. The precipitated protein was collected with a centrifuge (20,000g for 15 min). The obtained precipitate was suspended in 250 μL of dialysis buffer (50 mM potassium phosphate, 50 mM KCl, 1 mM EDTA, pH 6.8), and then dialyzed against the same buffer for 12 h.

Overexpression and purification of ubiquitin and RbpA1 in E. coli cells

Overexpression and purification of ubiquitin and RbpA1 were achieved as described previously (Sakamoto et al. 1999; Morita et al. 2000). Briefly, pET-24a and pET-21d plasmids harboring yeast *ubiquitin* and *RbpA1*, respectively, were transferred to *E. coli* BL21(DE3) cells. These cells were cultured at 37°C in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ and the protein expression was induced with IPTG. Cells were collected and disrupted with sonication. After centrifugation, different processes were achieved for further purification.

For yeast ubiquitin, the supernatant was incubated at 85°C for 5 min and then chilled on ice. After a further centrifugation step, the supernatant was further purified with ion-exchange column chromatography (SP Sepharose FF column; Pharmacia) and gel-filtration chromatography (HiLoad 26/60 Superdex 75 pg column; Pharmacia).

For RbpA1, the supernatant was purified with ion-exchange and gel-filtration column chromatography only (Q Sepharose FF column, and HiLoad 26/60 Superdex 75 pg column; Pharmacia).

The purified proteins were concentrated and applied to the NMR study.

Measurement of HSQC spectra

D_2O was added to the concentrated protein solutions (final concentration, 10%). HSQC spectra of these proteins were obtained with a DMX 500 (Bruker) FT-NMR spectrometer. The data were processed using NMRPipe (Delaglio et al. 1995) on a Linux work-

station. The ^1H and ^{15}N chemical shifts were referenced according to the method of Wishart et al. (1995).

Acknowledgments

This work was supported by a grant from National Project on Protein Structural and Functional Analyses.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Blaschke, U.K., Silberstein, J., and Muir, T.W. 2000. Protein engineering by expressed protein ligation. *Methods Enzymol.* **328**: 478–496.
- Chrnyk, B.A., Evans, J., Lillquist, J., Young, P., and Wetzel, R. 1993. Inclusion body formation and protein stability in sequence variants of interleukin-1 β . *J. Biol. Chem.* **268**: 18053–18061.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMR Pipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277–293.
- Golf, S.A. and Goldberg, A.L. 1987. An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in *Escherichia coli*. *J. Biol. Chem.* **262**: 4508–4515.
- Kurland, C.G. 1982. Translational accuracy in vitro. *Cell* **28**: 201–202.
- Madin, K., Sawasaki, T., Ogasawara, T., and Endo, Y. 2000. A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: Plants apparently contain a suicide system directed at ribosomes. *Proc. Natl. Acad. Sci.* **97**: 559–564.
- Morita, E.H., Murakami, T., Uegaki, K., Yamazaki, T., Sato, N., Kyogoku, Y., and Hayashi, H. 2000. NMR backbone assignments of the cold-regulated RNA-binding protein, RbpA1, in the cyanobacterium, *Anabaena variabilis* M3. *J. Biomol. NMR* **17**: 351–352.
- Ogasawara, T., Sawasaki, T., Morishita, R., Ozawa, A., Madin, K., and Endo, Y. 1999. A new class of enzyme acting on damaged ribosomes: Ribosomal RNA apurinic site specific lyase found in wheat germ. *EMBO J.* **18**: 6522–6531.
- Pavlov, M.Y. and Ehrenberg, M. 1996. Rate of translation of natural mRNAs in an optimized in vitro system. *Arch. Biochem. Biophys.* **328**: 9–16.
- Roberts, B.E. and Paterson, B.M. 1973. Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proc. Natl. Acad. Sci.* **70**: 2330–2334.
- Sakamoto, T., Tanaka, T., Ito, Y., Rajesh, S., Iwamoto-Sugai, M., Kodera, Y., Tsuchida, N., Shibata, T., and Kohno, T. 1999. An NMR analysis of ubiquitin recognition by yeast ubiquitin hydrolase: Evidence for novel substrate recognition by a cysteine protease. *Biochemistry* **38**: 11634–11642.
- Sato, N. 1995. A family of cold-regulated RNA-binding protein genes in the cyanobacterium *Anabaena variabilis* M3. *Nucleic Acid Res.* **23**: 2161–2167.
- Sawasaki, T., Ogasawara, T., Morishita, R., and Endo, Y. 2002. A cell-free protein synthesis system for high throughput proteomics. *Proc. Natl. Acad. Sci.* **99**: 14652–14657.
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. 2001. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**: 751–755.
- Spirin, A.S., Baranov, V.I., Ryabova, L.A., Ovodov, S.Y., and Alakhov, Y.B. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **25**: 1162–1164.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. 1995. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**: 135–140.