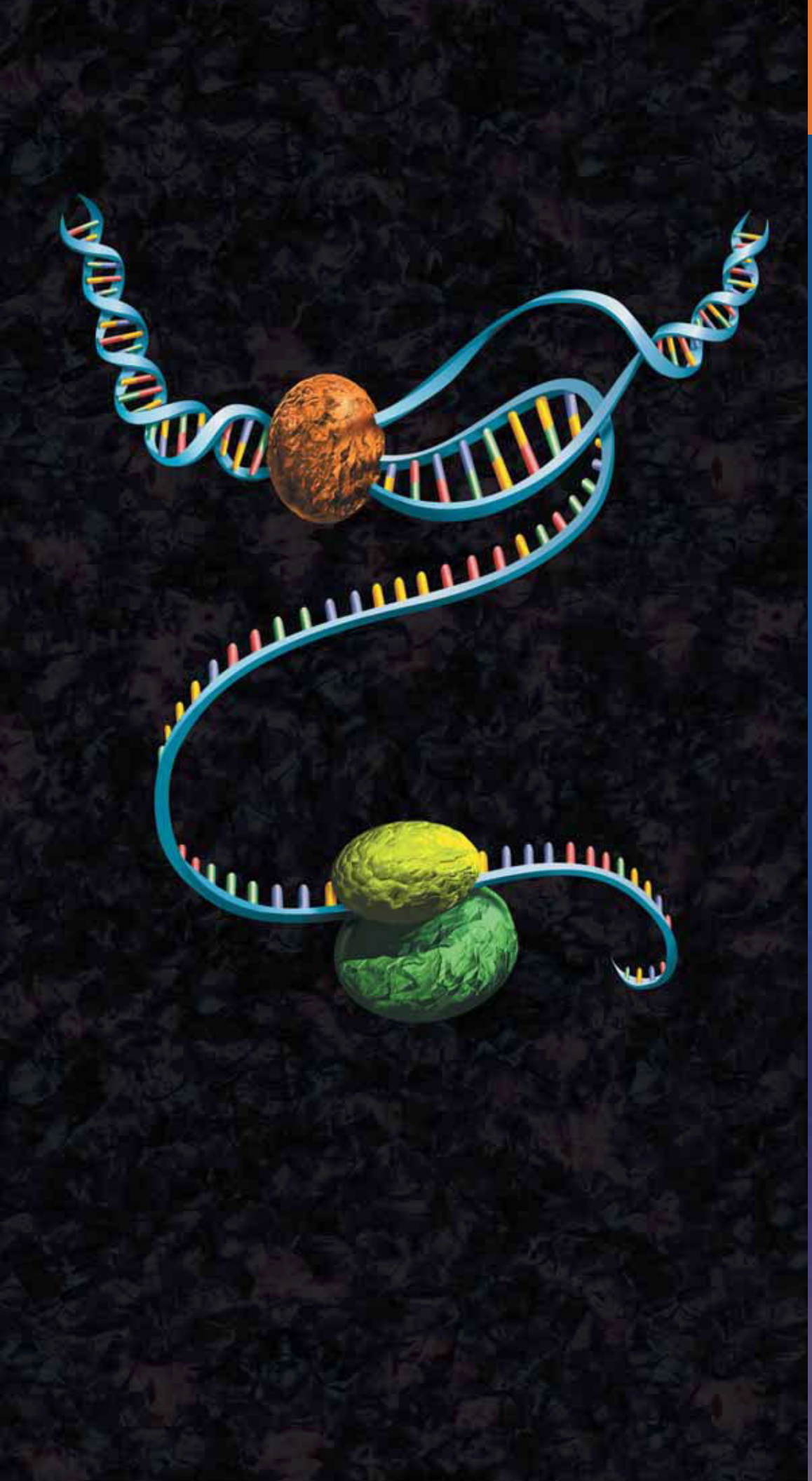

OVERVIEW OF IN VITRO EXPRESSION

CHAPTER 1

About the Image: *In this illustration of transcription and translation, the DNA double helix separates at the RNA polymerase molecule with an mRNA strand extending downward from the RNA polymerase. Towards the 5´-end of the mRNA the two subunits of a ribosome assemble to translate the RNA message into a polypeptide chain, which will fold into a protein molecule. Throughout this In Vitro Expression Guide various and recent applications of translation and protein expression are discussed.*

Promega

www.promega.com

Chapter One: Overview

Contents

We wish to thank Wiley-Liss, a subsidiary of John Wiley & Sons, Inc., for their permission to reproduce much of the material in the chapter by Jagus, R. and Beckler, G. (1998) Overview of eukaryotic in vitro translation and expression systems, *Current Protocols in Cell Biology*, 11.1.1–11.1.13, in this guide.

Introduction

The use of cell-free systems for the in vitro expression of proteins is a rapidly growing area, with applications in basic research, molecular diagnostics and high-throughput screening. In vitro expression encompasses two general strategies. The first is to use isolated RNA synthesized in vivo or in vitro as a template for the translation reaction (e.g., using Promega's Rabbit Reticulocyte Lysate[a,b,c] (Cat.# L4151) or Wheat Germ Extract (Cat.# L4380) Systems). The second is to use a coupled transcription/translation system in which DNA is used as a template (e.g., Promega's TnT® [a,b,c,d,e] and *E. coli* S30 Extract[a,b] Systems). This DNA may be either a gene cloned into a plasmid vector (cDNA) or a PCR[f]-generated template.

In vitro expression is a rapidly growing and constantly evolving field. This guide is intended to provide a general overview of the technology as presented in recent scientific publications.

Applications

Genetic Verification and Detection

Open Reading Frame (ORF) Expression. Probably still the most common use of in vitro expression is simply to decode a nucleic acid to determine if a gene or ORF is present. In DNA, an ORF consists of an initiation codon (usually ATG, but can also be GTG), followed by a sequence of nucleotides coding for amino acids and ending with a termination codon (TAA, TAG or TGA).

For example, a prokaryotic DNA (or eukaryotic cDNA) clone can be analyzed for the presence of ORFs using an in vitro expression system. Alternatively, in vitro expression can be used to verify an ORF predicted by DNA sequencing. In each case, the resulting proteins are characterized, and the protein size and structure are correlated to the size and sequence of the gene.

Specific uses include analysis of viral RNA genomes in which eukaryotic translation systems are used to determine the number and function of the viral genes (1), and study of differential protein expression using total cellular mRNA or polysomes from different tissues.

Cloned cDNA Expression. Cloned cDNAs positioned behind a phage promoter (e.g., T7, T3 or SP6) are commonly used for generating gene-specific mRNA used to program translation reactions (2). Run-off 5′-capped or uncapped mRNA can be produced in vitro and added to translation extracts. Alternatively, these DNA constructs can be used directly in eukaryotic coupled transcription/translation reactions.

Functional Analyses

Enzymatic Activity Analysis. Many proteins expressed using in vitro systems are correctly folded and processed and display normal in vivo enzymatic activity. If the extract system itself lacks (or has low levels of) the enzymatic activity of the expressed protein, the resulting translation reaction can be assayed directly without protein purification. Another advantage of in vitro expression is the ability to add exogenous factors to study enzymatic activity, potentially eliminating the need for transfection studies.

In one example, expression of adenylate cyclase (ACIV; 110kDa) in a TNT® System produced a protein with the same specific enzymatic activity as ACIV produced from a baculovirus expression system (3). In another example, aromatase produced in a TNT® System reaction supplemented with canine pancreatic microsomes and recombinant cytochrome P_{450} reductase resulted in an active enzyme (4).

Mutation Analysis. Upon cloning a gene, a number of studies are undertaken to discern the function of a gene product. A first step is often the introduction of a mutation into the gene to examine the effect on the expressed protein. Methods of mutagenesis include: i) serial deletion mutation analysis by progressive truncation of the 5′- or 3′-ends of the gene or by using *Bal* 31 digestion from an internal single-cut restriction site; and ii) site-directed point mutation analysis (e.g., using Promega's GeneEditor™ System(g)). Both methods can be used to identify functionally active domains or residues.

Post-Translational Modification Analysis. Post-translational modification of the protein, such as proteolytic cleavage or the addition of sugars, lipids, phosphate or adenyl groups, is often required for functional activity. Each in vitro expression system has its own endogenous post-translational modification activities. For example, various phosphorylation, adenylation, myristoylation, farnesylation, isoprenylation and proteolytic activities have been observed using rabbit reticulocyte lysate. Addition of microsomal membranes allows the study of glycosylation, methylation and removal of signal sequences.

Because not all the differences between the various in vitro expression systems are known, it may be desirable to try both reticulocyte lysate and wheat germ extract (and in some cases, *E. coli* S30 extract) to determine which system can produce a functional gene product with the "correct" post-translational modifications. In addition, cellular extracts or different microsomal membrane sources (e.g., *Xenopus* egg extracts) can be added to provide additional modifying activities.

Molecular Interaction Detection

Protein-Protein. Specific protein-protein interactions can be detected using in vitro expression methods. These interactions may include specific binding (such as antibody-antigen and ligand-receptor binding), macromolecular assembly, and formation of functional transcription complexes. In a common application, one protein partner is expressed in large amounts and purified from *E. coli* as a fusion protein. The other partner is expressed in an in vitro expression system as a labeled protein and used as a probe for detection of the interaction. Often, this technique is used to verify results from yeast two-hybrid experiments. A variety of biochemical analysis methods may be used to characterize the expressed proteins.

Protein-DNA and Protein-RNA. Putative DNA binding proteins, such as transcription factors, can be analyzed for their ability to bind to specific sequences on radiolabeled oligonucleotides. The binding is detected by an electrophoretic mobility shift assay (EMSA) in

References

1. Roberts, B.E. and Paterson, B.M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2330.
2. Melton, D.A. *et al.* (1984) *Nucl. Acids Res.* **12**, 7035.
3. Warner, D.R., Basi, N.S. and Rebois, R.V. (1995) *Anal. Biochem.* **232**, 31.
4. Pancharatnam, J. *et al.* (1996) *Molecular Zoology*, Ferraris, J.D. and Palumbi, S.R., eds., Wiley-Liss, New York.

References (continued)

5. Ebel, T. and Sippel, A. (1995) *Nucl. Acids Res.* **23,** 2076.
6. Milner, N., Mir, K.U., and Southern , E.M. (1997*) Nat. Biotechnol.* **15**, 537.
7. Lima, W.F. *et al.* (1997) *J. Biol. Chem.* **272**, 626.
8. Bayle, D., Weeks, D. and Sachs, G. (1997) *J. Biol. Chem.* **272**, 19697.
9. Bayle, D. *et al.* (1997) *J. Recept. Signal Transduct. Res.* **17**, 29.
10. Johnson, A.E. *et al.* (1976) *Biochemistry* **15**, 569.
11. Do, H. *et al.* (1996) *Cell* **85**, 369.
12. Hamman, B.D. *et al.* (1997) *Cell* **89**, 535.
13. Zhao, L. *et al.* (1999) *J. Biol Chem.* **274**, 14198.
14. Noren, C.J. *et al.* (1989) *Science* **244**, 182.
15. Rothschild, K.J. and Gite, S. (1999) *Curr. Opin. Biotech.* **10**, 64.
16. Frydman, J. *et al.* (1994) *Nature* **370**, 111.
17. Frydman, J. and Hartl, F.U. (1996) *Science* **272,** 1497.
18. Kolb, V.A., Makeyev, E.V. and Spirin, A.S. (1994) *EMBO J.* **13**, 3631.
19. Makeyev, E.V., Kolb, V.A. and Spirin, A.S. (1996) *FEBS Lett.* **378**, 166.
20. DiDonato, J.A. and Karin, M. (1993) *Promega Notes* **42**, 18.
21. Sakalian, M. *et al.* (1996) *J. Virol.* **70**, 3706.
22. Sonar, S. *et al.* (1993) *Biochemistry* **32**, 13777.
23. Roest, P.A. *et al.* (1993) *Hum. Mol. Genet.* **2**, 1719.

TO ORDER

Phone
1-800-356-9526

Fax
1-800-356-1970

Online
www.promega.com

Promega
www.promega.com

which greater retardation of the protein-DNA complex is observed when compared to unbound DNA. Usually the labeled DNA is added directly to the in vitro expression reaction. Researchers studying transcription factors such as NF-κB often use wheat germ extracts, as they do not contain endogenous mammalian transcription factors. A method has been reported to remove endogenous DNA-binding proteins from the reticulocyte system prior to the translation reaction (5).

DNA-RNA and RNA-RNA. Antisense DNA oligonucleotides can be useful for inhibiting expression at both the transcription and translation level. TnT® Systems have been used to rapidly screen oligonucleotides for those that best arrest translation (6,7).

Molecular Structure and Localization Analyses

Characterization of Membrane Association. In vitro expression systems have been successfully used to express integral membrane proteins. For example, expression of G-protein-coupled receptors in TnT® Systems supplemented with canine microsomal membranes results in the correct folding and insertion of transmembrane domains using the expressed signal anchor and stop transfer sequences (8,9)

Non-Natural Amino Acid Incorporation. Using technology originally developed in 1976, Johnson revolutionized in vitro translation by demonstrating that non-natural amino acids could be inserted into polypeptides using epsilon-modified, lysine-charged tRNAs (10). Further extension of this technology led to the incorporation of photoactivatible crosslinking or fluorescent groups into polypeptides, followed by monitoring of the molecular environment as the labeled peptides pass through the ribosome and enter the endoplasmic reticulum pore (11,12). Site-directed incorporation of a photoactivable crosslinker through a non-natural amino acid incorporation was used to capture protein interacting with different portions of a protein of interest (13). Other groups have developed site-specific methods that utilize an amber suppressor tRNA charged with any number of non-natural amino acids, including fluorescent, spin-label and isotopic groups (14).

For a review of tRNA-mediated protein engineering (TRAMPE) see reference 15.

Protein Folding and Chaperonin Interactions. In vitro expression is increasingly being used to understand the nature of sequential chaperonin interactions required for protein folding and localization. Researchers in this field have combined the advantages of in vitro expression with the power of instantaneous reporter gene product assays. The folding of polypeptides emerging from ribosomes has been analyzed using firefly luciferase as a model protein (16,17).

Real-Time Translation/Folding Assays. A novel approach has been developed using a wheat germ system in which the components for the luciferase enzymatic assay have been added directly to the translation reaction and monitored continuously in real time. Luciferase was shown to be fully folded and enzymatically active immediately upon release from the ribosome (18). However, no luciferase activity was observed while full-length luciferase remained attached to the ribosome as a peptidyl-tRNA, probably because the C-terminal portion of the enzyme is masked by the ribosome or ribosome-associated proteins. The investigators demonstrated that the ribosome-bound enzyme acquires enzymatic activity when its C-terminus is extended by at least 26 additional amino acid residues (19). The results demonstrate that the acquisition of the final native conformation by a nascent protein can occur as the protein is being synthesized and that folding does not require release of the protein from the ribosome.

Macromolecular Assembly. It is possible to express numerous gene products in one coupled transcription/translation reaction to form functional transcription factor complexes (20) or viral particles (21) that are identical to those formed in the host.

Molecular Structure Analysis. Understanding the function of integral membrane proteins is currently limited by the difficulty of producing crystals for use in X-ray diffraction studies. A method has been developed for probing conformational changes in membrane proteins using Fourier transform infrared-difference (FTIR) spectroscopy. In this method, natively folded polypeptides are expressed in vitro with a site-specific insertion of a single isotopic label through amber suppression (22). This method does not disrupt the protein structure as did earlier site-directed mutagenesis methods and should be applicable to a wide range of other proteins, including those involved in enzyme catalysis, ion transport and signal transduction.

Molecular Diagnostics

Protein Truncation Test. A growing application of coupled transcription/translation systems has been for diagnosis of genetic diseases, a DNA technology-dominated field. The protein truncation test (PTT), sometimes referred to as in vitro synthesized protein truncation (IVSP) assay, was first reported in 1993 as a rapid method for detecting translation-terminating mutations in the large gene responsible for Duchenne Muscular Dystrophy (23) and the Adenomatous Polyposis Coli *(APC)* gene responsible for a type of hereditary colon cancer (24). In these and other

diseases, such as hereditary breast cancer (24), 70–95% of the mutations that cause disease result in a truncated gene product.

The PTT involves first purifying genomic DNA or mRNA from the patient's blood or tissue. This is followed by either RT-PCR(f) or PCR(f) with the concurrent incorporation of a T7 promoter and optimal translation initiation sequence surrounding the desired start codon (25). Often when the source of mRNA is limiting, a second nested PCR amplification is required. Large exons are amplified from genomic DNA while smaller exons are amplified together from mRNA and the gene is segmented into overlapping amplified fragments. The amplified DNA is added directly to a coupled transcription/translation reaction and translation terminating mutations are detected as faster migrating bands after SDS-PAGE analysis. PTT has the advantage of enabling scans of large (2–3kb) DNA/RNA segments quickly. In addition, PTT detects only disease-causing mutations. This avoids the fruitless evaluation of polymorphisms. The recent introduction of the TNT® T7 Quick for PCR DNA System(c,d,e,g) facilitates PTT analysis.

High-Throughput Screening (26)

Screening for Viral-Specific Translation Inhibitory Compounds. Viruses contain a number of different genetic elements used for promoting viral expression at the expense of host mRNA translation. Several groups are currently developing screens using in vitro expression of gene constructs containing a viral element such as the 5′-UTR that can harbor an IRES (Internal Ribosome Entry Site) followed by a firefly or *Renilla* luciferase gene (27). Chemical or antibiotic libraries can be screened for specific translation-inhibiting effects. The viral element can be placed between the firefly and *Renilla* luciferase genes with translation of the first gene relying on normal cap-dependent initiation. Use of the two luciferase genes allows normalization of the second reporter behind the viral element. The efficacy of compounds can be assessed rapidly (<30 seconds) by assaying light output from both the reporter and the control luciferase (Figure 1, Chapter 5).

As a variation on this theme, RiboGene, Inc., has reported developing a high-throughput system for screening several hundred thousand compounds for the ability to diminish or block the required ribosomal frameshifting used during translation of the HIV gag-pol mRNA. This screen utilizes a reporter gene in which luciferase (and light) is produced only when the frameshift occurs.

Screening for Chaperonin-Inhibiting Drugs. The in vitro luciferase folding/chaperonin assay described earlier has been extended to understanding the role of heat shock factors, such as Hsp90. It is now understood that disruption of the folding pathways can result in proteolytic degradation. Several groups are currently using this information to ascertain the pharmacological activities of benzoquinone ansamycins, such as geldanamycin (28,29). These potentially medically important compounds were first identified as interesting because of their ability to inhibit tyrosine kinase activity. This ability appears to be due to their interaction with Hsp90, which prevents the correct folding of tyrosine kinases and is followed by their proteolytic degradation. Other potentially important drugs affecting protein folding through inhibition of chaperonin function could be identified using this approach.

Identification of Novel Orphan Receptors. The binding of ligands to in vitro synthesized receptors can be an important aspect of identifying new receptors. For example, in a search for novel "orphan" nuclear receptors and ligands, a novel estrogen receptor was cloned and characterized (30). Saturation ligand-binding and ligand-competition assays of the in vitro expressed clone allowed this novel receptor to be distinguished from a previously cloned receptor.

Functional Genomics

In Vitro Expression Cloning (IVEC). In this procedure, an oligo(dT)-primed cDNA library is constructed in a high copy expression plasmid containing a T3, T7 or SP6 promoter. The plasmid library is then transformed into *E. coli*, and approximately 10^5 independent transformants are plated on selective media. The bacterial colonies are grown to a specific size (e.g., 1mm in diameter), collected and pooled (50–100 clones per pool). Purified plasmid DNA from these pools is directly added to a small-scale (e.g., 10µl) coupled transcription/translation reaction, where it is used as a template in the presence of [^{35}S]methionine (31). Depending upon the number of the full-length cDNA clones in the library, approximately 30–50 proteins can be produced in a single reaction. Proteins can be assayed for any number of activities, including phosphorylation, proteolysis or cleavage. Positive pools are subdivided until the single cDNA that encodes the protein of interest is isolated.

Ribosomal Display for Cell-Free Protein Evolution. In this procedure, a cell-free system is used to transcribe a DNA library, translate the mRNA pools and, using a variety of techniques, the proteins and the encoding mRNAs are retained, still attached to the ribosomes. The protein-mRNA-ribosome complexes are screened for binding to a target, and the retained mRNA is

References (continued)

24. Powell, S.M. *et al.* (1993) *N. Engl. J. Med.* **329**, 1982.
25. Hogervorst, F.B. *et al.* (1995) *Nat. Genet.* **10**, 208.
26. Jagus, R. and Beckler, G.S. (1998). Overview of eukaryotic in vitro translation and expression systems. *Current Protocols in Cell Biology* 11.1.1–11.1.13. Copyright © 1998 by John Wiley & Sons, Inc. Reproduced by permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.
27. Grentzmann, G. *et al.* (1998) *RNA* **4**, 479.
28. Schneider, C. *et al.* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 14536.
29. Thulasiraman, V. and Matts, R.L. (1996) *Biochemistry* **35**, 13443.
30. Kuiper, G.G. *et al.* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 5925.
31. *TNT® Coupled Reticulocyte Lysate Systems Technical Bulletin* #TB126, Promega Corporation.

References (continued)

32. Mattheakis, L.C., Bhatt, R.R., and Dower, W.J. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9022.
33. Hanes, J. and Plückthun, A. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 4937.
34. He, B., Gross, M., and Roizman, B. (1996) *Proc. Natl. Acad. Sci. USA* **94**, 843.
35. Joyce, G.F. (1993) *Pure Appl. Chem.* **65**, 1205.

amplified using RT-PCR® with the resulting DNA used for another round of selection. Initially, a prokaryotic *E. coli* coupled transcription/translation system was used to generate large libraries of peptides for receptor ligand screening (32). Later improvements allowed folding of whole proteins into their native structure while still attached to the ribosome (33). The first eukaryotic application used a coupled rabbit reticulocyte system to study antibody-ribosome-mRNA (ARM) complexes, allowing for rapid selection and monitoring of antibody combining site evolution (34).

A cell-free system has been developed for performing evolution studies in which RNA amplification and the coupled reaction can be performed simultaneously at a given temperature (35). After unsuccessful attempts using wheat germ extracts and coupled *E. coli* systems, investigators were able to combine the reactions using a rabbit reticulocyte coupled system. By exerting selective pressure on functional protein products necessary for RNA amplification, this system can be used for performing laboratory "evolution."

Preparative Synthesis

Large-Scale Protein Expression and Purification. Cell-free expression systems are often preferred over in vivo or native systems, because they can be used for the expression of toxic, proteolytically sensitive or unstable proteins. In addition, in vitro systems provide the ability to incorporate non-natural amino acids containing photoactivatable, fluorescent or biotin residues. Typically, in vitro systems produce nanogram amounts of proteins per 50µl reaction; however, preparative scale methods have been developed recently that may yield milligram quantities per milliliter of reaction mixture.

Guide Organization

This chapter provides a general overview of many of the major applications of in vitro expression systems. The remainder of the guide focuses on several of these applications, seeking to provide more detailed information on commonly used methodologies for in vitro expression technology.