METHODS IN MOLECULAR BIOLOGY™

High Throughput Protein Expression and Purification

Methods and Protocols

Edited by

Sharon A. Doyle

GlaxoSmithKline, Walnut Creek CA, USA

Humana Press

Editor
Sharon A. Doyle
GlaxoSmithKline
Walnut Creek CA
USA

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

ISBN: 978-1-58829-879-9 e-ISBN: 978-1-59745-196-3
ISSN: 1064-3745 e-ISSN: 1940-6029
DOI: 10.1007/978-1-59745-196-3

Library of Congress Control Number: 2008931181

Cover illustration:

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

Advances in genome sequencing have changed the scope of scientific inquiry for the biological sciences. Entire genome sequences can now be determined rapidly, providing researchers with the raw data that encode all proteins produced by an organism. One can now devise experiments to discover the structures, functions, and interactions of these proteins on a cellular or organismal level that previously was not possible. Scientists in a wide range of fields including structural biology, functional genomics, and drug discovery are poised to benefit greatly from these advances.

This exciting opportunity has brought with it new challenges because proteins do not behave in a uniform and predictable manner as do their DNA counterparts. There is no universal method to isolate a protein from its expression system in its native form, due to potentially large variations in size, charge, shape, and external chemical moieties. In addition, the conditions required for proper folding and stability, for example, vary greatly among proteins.

Yet for experiments to yield large numbers of proteins, parallel processing is a necessity. Biochemists, instrumentation engineers, and bioinformaticists have recognized this challenge and over the last few years have made tremendous progress on developing technologies to enable high-throughput protein production. This includes not only instrumentation to handle large numbers of samples in parallel, but also strategies that create and exploit common features to enable simplified cloning, stable expression, and purification of many types of proteins.

This comprehensive volume presents current methodologies including various high-throughput cloning schemes, protein expression analysis, and production protocols. Methods are described that utilize *E. coli*, insect, and mammalian cells, as well as cell-free systems for the production of a wide variety of proteins, including glycoproteins and membrane proteins. This volume provides details of the most successful protocols currently in use by leading scientists, presented in a manner that we hope will be useful to those with training in protein biochemistry, as well as to those new to laboratory-based protein biochemistry who plan to apply these protocols to related fields.

Walnut Creek, CA — Sharon Doyle

Contributors

JAMES M. ABDULLAH • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

DAVID ALDERTON • *Oxford Protein Production Facility, Welcome Trust Centre for Human Genetics, Oxford, UK*

ERIN S. ARROYO • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

DOUGLAS A. AUSTEN • *Gene Expression, Vertex Pharmaceuticals, Cambridge, MA, USA*

BRIAN P. AUSTIN • *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA*

MAHMOUD REZA BANKI • *Department of Chemical Engineering, Princeton University, Princeton, NJ, USA*

GRAHAM BENCH • *Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

NICK S. BERROW • *Oxford Protein Production Facility, Welcome Trust Centre for Human Genetics, Oxford, UK*

CRAIG D. BLANCHETTE • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

PAUL G. BLOMMEL • *Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA*

MIRJAM BUCHS • *Biologics Center, Novartis Institutes for Biomedical Research, Basel, Switzerland*

JENNY A. CAPPUCCIO • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

CHACKO S. CHAKIATH • *Protein Expression Laboratory, SAIC-Frederick, Inc., National Cancer Institute at Frederick, MD, USA*

STEPHEN P. CHAMBERS • *Gene Expression, Vertex Pharmaceuticals, Cambridge, MA, USA*

DEB K. CHATTERJEE • *Protein Expression Laboratory, SAIC-Frederick, Inc., National Cancer Institute at Frederick, MD, USA*

SCOTT CHERRY • *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA*

BRETT A.CHROMY • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

MATTHEW A. COLEMAN • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

FRANK R. COLLART • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

MICHELE H. CORZETT • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

Mark I. Donnelly • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

William H. Eschenfeldt • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

Said Eshaghi • *Division of Biophysics, Department of Medical Biochemistry and Biophysics, Karolinska Instiutet, Stockholm, Sweden*

Dominic Esposito • *Protein Expression Laboratory, National Cancer Institute, SAIC-Frederick, Frederick, MD, USA*

Julia Fletcher • *Cloning and Protein Expression, R&D, Invitrogen Corporation, Carlsbad, CA, USA*

Brian G. Fox • *Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA*

John R. Fulghum • *Gene Expression, Vertex Pharmaceuticals, Cambridge, MA, USA*

Leslie A. Garvey • *Protein Expression Laboratory, SAIC-Frederick, Inc., National Cancer Institute at Frederick, MD, USA*

Sabine Geisse • *Biologics Center, Novartis Institutes for Biomedical Research, Basel, Switzerland*

Alison R. Gillies • *Department of Chemical Engineering, Princeton University, Princeton, NJ, USA*

Paul T. Henderson • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

Angela K. Hinz • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

Paul D. Hoeprich Jr. • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

Ian Hunt • *Protein Structure Unit, Centre for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Cambridge, MA, USA*

Andrzej Joachimiak • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

Federico Katzen • *Cloning and Protein Expression, R&D, Invitrogen Corporation, Carlsbad, CA, USA*

Brian K. Kay • *Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL, USA*

Vladimir Kery • *Biomimetic Therapeutics, Franklin, TN, USA*

Ernie Kim • *Protein Structure Unit, Centre for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Cambridge, MA, USA*

Heath Klock • *The Genomics Institute of the Novartis Research Foundation, San Diego, CA, USA*

Jim Koehn • *Protein Structure Unit, Centre for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Cambridge, MA, USA*

Wieslaw Kudlicki • *Cloning and Protein Expression, R&D, Invitrogen Corporation, Carlsbad, CA, USA*

Edward A. Kuhn • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

RICHARD J. LAW • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

SCOTT A. LESLEY • *The Genomics Institute of the Novartis Research Foundation, San Diego, CA, USA*

CHIANN-TSO LIN • *Molecular Biosciences, Pacific Northwest National Laboratory, Richland, WA, USA*

FAN LU • *Gene Expression, Vertex Pharmaceuticals, Cambridge, MA, USA*

MARION MAHNKE • *Biologics Center, Novartis Institutes for Biomedical Research, Basel, Switzerland*

PETER A. MARTIN • *Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA*

CYNTHIA SANVILLE MILLARD • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

PRISCILLA A. MOORE • *Molecular Biosciences, Pacific Northwest National Laboratory. Richland, WA, USA*

SREEDEVI NALLAMSETTY • *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA*

JOANNE E. NETTLESHIP • *Oxford Protein Production Facility, Welcome Trust Centre for Human Genetics, Oxford, UK*

RAYMOND J. OWENS • *Oxford Protein Production Facility, Welcome Trust Centre for Human Genetics, Oxford, UK*

JOSEPH B. PESAVENTO • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

TODD PETERSON • *Cloning and Protein Expression, R&D, Invitrogen Corporation, Carlsbad, CA, USA*

YANN POULIQUEN • *Biologics Center, Novartis Institutes for Biomedical Research, Basel, Switzerland*

NAHID RAHMAN-HUQ • *Oxford Protein Production Facility, Welcome Trust Centre for Human Genetics, Oxford, UK*

MICHAEL SACHS • *Protein Structure Unit, Centre for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Cambridge, MA, USA*

KORY D. SEDER • *Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA*

BRENT W. SEGELKE • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

KALAVATHY SITARAMAN • *Protein Expression Laboratory, SAIC-Frederick, Inc., National Cancer Institute at Frederick, MD, USA*

LUCY STOLS • *Biosciences Division, Argonne National Laboratory, Lemont, IL, USA*

TODD A. SULCHEK • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

SUSANNE E. SWALLEY • *Gene Expression, Vertex Pharmaceuticals, Cambridge, MA, USA*

SANG THAI • *Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL, USA*

JOSEPH E. TROPEA • *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA*

VERONICA V. VOLGINA • *Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL, USA*

VICKI L. WALSOWRTH • *Biosciences and Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA, USA*

DAVID S. WAUGH • *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA*

DAVID W. WOOD • *Department of Chemical Engineering, Princeton University, Princeton, NJ, USA*

RUSSELL L. WROBEL • *Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA*

Contents

Chapter 1

High-Throughput Protein Production (HTPP): A Review of Enabling Technologies to Expedite Protein Production

Jim Koehn and Ian Hunt

Summary

Recombinant protein production plays a crucial role in the drug discovery process, contributing to several key stages of the pathway. These include exploratory research, target validation, high-throughput screening (HTS), selectivity screens, and structural biology studies. Therefore the quick and rapid production of high-quality recombinant proteins is a critical component of the successful development of therapeutic small molecule inhibitors. This chapter will therefore attempt to provide an overview of some of the current "best-in-class" cloning, expression, and purification strategies currently available that enhance protein production capabilities and enable greater throughput. As such the chapter should also enable a reader with limited understanding of the high-throughput protein production (HTPP) process with the necessary information to set up and equip a laboratory for multiparallel protein production.

Key words: Deep-well block protein expression; Miniaturized protein purification; High throughput

1.1. Introduction

Over the last few years a number of technologies have been developed to expedite the production of recombinant proteins for therapeutic studies. These include the use of rapid cloning systems, miniaturization of cell growth conditions, and a variety of innovative automation systems for expression and purification of recombinant proteins. A quick Web search easily identifies several companies offering complete high-throughput protein production (HTPP) systems. However true HTPP is the preserve of only a few laboratories and/or consortia worldwide whose focus is the rapid generation of many hundreds of proteins simultaneously. What technologies are available to those labs that require only a modest, but significant improvement in throughput?

Sharon A. Doyle (ed.), *Methods in Molecular Biology: High Throughput Protein Expression and Purification, vol. 498*

Book doi: 10.1007/978-1-59745-196-3

This question is probably relevant to most labs where target prioritization and interest necessitate they work on specific protein targets that can often be difficult to express. This review will therefore attempt to provide an overview of some of the current "best–in-class" cloning, expression, and purification strategies currently available. Specifically, the review will focus on some of the developments in *E. coli*- and baculovirus-mediated protein expression systems that have enhanced protein production within the context of the drug discovery environment.

1.2. High-Throughput Cloning Methods

One of the standard procedures when setting out to express a recombinant protein is to screen a series of constructs to identify the most viable protein for the generation of sufficient soluble material for downstream purposes. This may include expressing full-length proteins or perhaps specific domains or chimeric proteins. A series of fusion partners may also be investigated for their effects on driving enhanced expression or their capacity to capture and purify the target protein quickly with minimal impurities. Using traditional cloning methodologies, the generation of the many possible combinations and their analysis in different expression systems would be so labor intensive and time consuming as to make a parallel strategy of expression screening impractical. However, over the past few years many of the limitations relating to the generation of multiple expression plasmids (and constructs) have been addressed by a number of elegant recombinatorial cloning systems that enable the rapid cloning of potentially hundreds of genes and constructs simultaneously (*see* **Table 1.1**). By far the most popular are Gateway® (Invitrogen) and LIC (Novagen, EMD Biosciences), although the former does result in the expression of extraneous coding regions that may be deleterious to downstream applications (e.g., crystallography). Moreover, the long-term cost in adopting these technologies can be substantial. Therefore, several reports have detailed low-cost alternatives that offer comparable capabilities. Most notable are those described by Klock et al. *(1)* and Benoit et al. *(2)*. In both these cases, the resulting expression plasmids contain no extraneous coding regions and thus offer significant advantages to those approaches which are commercially available. In addition, both systems are easily adapted to liquid-handling workstations.

Irrespective of the method employed for the construction of the expression plasmid(s), propagation of sufficient recombinant DNA for expression studies necessitates plasmid selection and purification. Nowadays (even for large numbers of constructs)

Table 1.1
Comparison of rapid direction cloning strategies

	Cloning system					
	Traditional	**TOPO**	**Gateway**	**Creator**	**In-fusion**	**LIC**
Commercial supplier		Invitrogen http://www.invitrogen.com	Invitrogen	Clontech http://www.clontech.com	Clontech	Novagen http://www.emdbiosciences.com
Mechanism of gene insertion	Classic sticky end ligation using T4 DNA ligase	TOPO cloning technology	Lambda bacteriophage	Cre-loxP		T4 DNA polymerase
Efficiency	Varies	Varies; insertion of large DNA fragments (1,500 bp) can be problematic	>90%	>90%	>90%	>90%
Potential use in multiparallel expression strategies	Limited	Yes	Yes	Yes	Yes	Yes
Incorporation of additional amino acids	Yes, typically the addition of 2 residues corresponding to restriction sites	No	Addition of 8 amino acid residues	Introduction of Cre-LoxP regions	No	Yes, addition of 5 amino acid residue. Can be removed by enterokinase cleavage
Comments		Potentially any expression vector can be adapted. Conversion to TOPO systems is however currently costly			Following ligation requires transfer to Creator. Cloning vehicle only	

this is a relatively straightforward process; following single-plate transformation (*see* **Note 1**), positive clones can be selected, propagated, and purified using a variety of standard commercial high-throughput kits that are easily adapted to liquid-handling workstations (QIAwell Ultra Plasmid BioRobot Kit, Qiagen; Wizard SV 96 Plasmid DNA Purification Kit, Promega; Perfect Prep Plasmid 96 Vac Direct Bind Kit, Eppendorf and Nucleospin Robot-96 Plasmid Kit, Macherey-Nagel). Most of these commercial kits can either be used in a standalone format with minimal equipment requirements (96-single tube well magnet, Qiagen Cat no. 36915 or vacuum manifold), or adapted to a variety of liquid-handling systems (*see* **Section 1.6** for further details). For each specific robot, it is recommended to refer to the kit manufacturer's Web site for detailed information on plate/machine compatibility since this can vary markedly.

1.3. Expression Systems

1.3.1. E. coli

Bacterial expression is the most commonly employed expression vehicle for the production of recombinant proteins. It is relatively simple to manipulate, inexpensive to culture, and the amount of time necessary to generate a recombinant protein is relatively short. Recombinant expression of proteins is normally achieved through the induction of a strong promoter system. Several are commonly used including T7, lambda Pl, and *ara*B. Perhaps the most popular is the T7-based pET expression plasmids (commercially available from Novagen). However, while this system leads to the generation of large amounts of mRNA and concomitant protein expression, the high levels of mRNA can cause ribosome destruction and cell death. Furthermore, leaky expression of T7 RNA polymerase may result in plasmid or expression instability. Use of the lac operator and T7 lysozyme (pLysS) can, however, provide an extra level of repression. A more complete review of *E. coli* expression vectors can be found elsewhere *(3–5)*.

1.3.1.1. HT E. coli Autoinduction

A very useful tool in the expression testing of multiple proteins and conditions, autoinduction was first described by Studier *(6)*. Autoinduction refers to bacterial cultures grown in media containing specific components that after an initial period of tightly regulated, uninduced growth, automatically induces target protein expression without IPTG. Typically, autoinduced expression produces a greater proportion of soluble target protein than does IPTG-induced expression. This method can provide higher protein yields and greater convenience compared to standard IPTG induction,

in particular, when working with large numbers of parameters and or constructs. Several autoinduction reagents are sold commercially including Overnight Express™ autoinduction from Novagen (http://www.emdbiosciences.com/novagen).

1.3.1.2. Expression Strains

Many *E. coli* strains optimized for protein expression purposes are commercially available from suppliers such as Invitrogen, Novagen, and Stratagene. The Origami™ and Origami B™ strains (Novagen) have been developed to express proteins that require disulfide bonds to achieve their active, correctly folded conformation. These strains with mutations in both the thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes greatly enhance disulfide bond formation in the cytoplasmic space. Furthermore, strains such as BL21 CodonPlus™ (Stratagene) and Rosetta-2™ derived from BL21 (Novagen) have been reported to enhance expression of gene sequences that contain rarely used *E. coli* codons in the expression of heterologous protein in *E. coli*. Strains that have been engineered to express the *Oleispire anrarctica* chaperones Cpn10 and Cpn60 have also been shown to enhance expression. Commercially available as Artic-Express™ (Stratagene), this strain is particularly useful in conferring improved protein expression levels at lower temperatures, potentially increasing the yield of active, soluble recombinant proteins. Many of these strains are sold in 8-well strips and 96-well plate formats, allowing convenient transfer of protocols to HT formats using liquid-handling workstations.

1.3.2. Baculovirus

A detailed review of the different baculovirus systems and their application to HT protein expression is covered elsewhere *(7)*; however, for completeness we briefly review some of the developments that have enhanced (at least in part) the streamlining and speed of baculovirus-mediated insect cell expression. In this regard it is noteworthy that while the Bac-to-Bac® system (Invitrogen) is perhaps the most popular insect cell expression system, it is one of the least amenable to HT applications. The large quantities of recombinant bacmid that is required for transfection make the system very time consuming and cumbersome, especially when working with 24–96 different constructs. Conversely transient insect cell systems (Insect Direct™, Novagen) and some of the newer baculovirus-mediated systems (FlashBAC™, NextGen Sciences and BacMagic™, Novagen) that do not require bacmid propagation are actually more amenable to HTPP and automation. Therefore, it will be interesting to see protein science evolve toward higher throughput, if we see a switch in the popularity of the respective systems.

1.3.2.1. HT Bacmid Propagation (Bac-to-Bac Only)

The generation of recombinant bacmid can be performed by the use of several commercial kits and a detailed protocol is described elsewhere in this book. Both the R.E.A.L Prep 96 Kit (Qiagen)

and PerfectPrep™ BAC 96 (Eppendorf/Brinkmann) kits are amenable for HT purification of bacmid DNA by both manual and liquid-handling strategies.

1.3.2.2. HT Transfection

The use of the Amaxa Nucleotransfection system (*see* http://Amaxa.com) provides a very fast and highly efficient method of performing large numbers of suspension-based transfections of mammalian and insect cells. The system can simultaneously transfect 8–96 constructs in approximately 5 min into 50–100 µL suspension culture volumes. Since the process is extremely efficient and can be conducted in suspension-based cultures, protein expression studies can be completed in a relatively short period of time (following a single round of viral application). This represents a significant advantage over traditional (manual) approaches in which transfection can take 2–3 h to complete and which are extremely laborious when working with 24 or more different constructs. Taken together, the system represents a significant breakthrough in streamlining the HT insect cell expression process. Moreover, coupled with the use of transient or non-bacmid based baculovirus expression strategies, this approach offers a very powerful system for insect-driven HTPP.

1.3.3. Other Expression Systems

Several other systems are also used extensively in the expression of recombinant proteins for structural, functional, and high-throughput assays. However, for the sake of brevity this review has focused on those expression systems most routinely encountered in HTPP laboratories.

1.3.3.1. Mammalian Cell Expression

The expression of heterologous proteins in a mammalian background offers many clear advantages to their generation in *E. coli* or insect cells, including correct post-translational modification and folding. However, while the use of mammalian cells such as CHO or HEK293 is well documented, the process of creating stable mammalian cell lines can often be laborious and time consuming *(8)*. Transient expression systems that utilize suspension cell lines therefore may provide a viable alternative, producing high quantities of recombinant protein in a very short period of time. Of particular interest is the use of HEK293-EBNA cells for rapid transient expression studies *(9–11)*. The system utilizes episomally replicating plasmids featuring the Epstein–Barr virus (EBV) *oriP* driven by EBNA-1 protein generated from a gene integrated into the HEK293 genome. In most cases transient protein expression is driven by the strong CMV promoter. Furthermore, the system is highly amenable to automation, and many groups have begun to adapt the system to miniaturized high-throughput strategies in a similar fashion to those described for *E. coli* and baculovirus incorporating FLAG and polyhistidine tags to facilitate rapid HT-purification and quantitation.

1.3.3.2. Cell-Free Expression

Traditional cell-free expression systems, while providing an attractive "quick" route to the production of proteins, have always been marred by low expression levels. However, several improvements have been made that now enable expression of 5–10 mg/mL. Changes include the optimization of lysate composition, introduction of semi-continuous reactions, and energy regeneration systems *(12–15)*. The use of a eukaryotic-based approach coupled with the rapid, cell-free utility offered by the system has a number of clear advantages and offers a viable alternate to traditional approaches. Indeed, some groups, notably RIKEN Structural Genomics Initiative, appear almost exclusively to generate recombinant protein in this way *(16, 17)*. Several commercial cell-free system are currently available and include the Rapid Translation System™ (RTS) platform (Roche), EasyXpress™ (Qiagen), and Expressway™ (Invitrogen), the former offering a complete scalable system for small-scale PCR-mediated screening to large-scale (10 mL) production in a variety of cell lysates (*E. coli* and wheat-germ bases). However, the expression of proteins that require SH-bond formation or membrane expression is not possible in either case. Furthermore, the commercial systems, while very quick and amenable to HT strategies, are very expensive and therefore are cost prohibitive for many laboratories.

1.4. Protein Expression and Purification Strategies

1.4.1. Miniaturized Protein Expression and Purification Screening

E. coli being a very robust organism is highly amenable for growth in 96- and 24-deep-well block formats (1–3 mL) using standard shaking incubators (*see* **Note 2**). Conversely, the growth of deep-well block insect cell cultures is somewhat more problematic. Indeed, while insect cells have been shown to grow successfully in 4–10 mL volumes using 24-deep-well blocks, reliable growth using 48- and 96-well blocks is difficult to achieve. Notwithstanding, by careful optimization of agitation speeds and humidity of either conventional shakers or those designed specifically for deep-well blocks, insect cells can be cultivated and used for transfection, viral amplification, and protein expression studies *(18–20)*.

Deep-well block expression requires a concomitant method of analyzing the resulting recombinant protein expression data. To this end, several groups have developed reporter assays to quickly assess the solubility of recombinant protein directly from the culture supernatant, thereby avoiding time-consuming and labor-intensive extraction and purification of the protein of interest *(21–24)*. However, in many cases additional information

regarding the expressed protein is required which necessitates the lysis, purification, and analysis of the recombinant protein. Several commercial lysis buffers are available and are listed. Some that do not require the harvesting of cellular debris by centrifugation have also been developed. Sold commercially as POP Culture™ (Novagen) and FastBreak™ (Promega) their use greatly facilitates the downstream analysis of protein expression, since they allow in situ analysis of cell lysates without the need for cellular clarification via centrifugation (*see* **Note 3**). This can be extremely advantageous when setting up an HTPP strategy on a liquid-handling workstation, since centrifugation steps can add significant time onto the process.

1.4.1.1. Affinity Tags for Solubility and HT Purification

The use of such a multiparallel expression strategy to conduct optimization and screening studies (whether conducted in a prokaryotic or eukaryotic host) necessitates the adoption of a generic purification approach that is quick, simple, and cheap to perform. In addition to offering a convenient one-step generic purification strategy, several fusion partners have also been shown to enhance solubility of the target protein. An excellent summary of the findings of various studies was presented by Braun and LaBaer *(25)*. Clearly, the use of many tags in a multiparallel expression strategy is impractical since the requirements for the different methods of purification and analysis would make for a logistically huge undertaking. Therefore typically most groups have their favorites based on experience [and success] and will preferentially use these in their work, often in concert with each other. Indeed, when used in tandem, affinity, solubility-enhancing, or reporter tags can provide very useful generic purification tools. For example, two affinity tags can be used in concert to produce high-purity protein via a two-step chromatography procedure, while GFP, S-tags, Trx, or NusA can be used in tandem with His6 and MBP to rapidly analyze and quantify expression levels and also allow downstream purification.

However, it is always important to remember that fusion tags (irrespective of size) can potentially interfere with folding, function, or crystallization capabilities. Conversely, there are many instances of proteins whose expression has been enhanced by the use of tags but which have become unfolded and precipitate following removal of the tag by enzymatic cleavage at an engineered recognition site adjacent the tag.

Deciding which to use is therefore difficult, since they all appear to offer various advantages over each other. Furthermore, some are far more amenable to adaptation to high-throughput purification strategies than others (most notably the His6 tag). A list of some of the commonly used tags that are amenable to rapid single-step purification of proteins that are commercially available is shown in **Table 1.2**. However as mentioned earlier,