1
Key and Criteria to the Selection of an Expression Platform

Gerd Gellissen, Alexander W.M. Strasser, and Manfred Suckow

The production of recombinant proteins has to follow an economic and qualitative rationale, which is dictated by the characteristics and the anticipated application of the compound produced. For the production of technical enzymes or food additives, gene technology must provide an approach which has to compete with the mass production of such compounds from traditional sources. As a consequence, production procedures have to be developed that employ highly efficient platforms and that lean on the use of inexpensive media components in fermentation processes. For the production of pharmaceuticals and other compounds that are considered for administration to humans, the rationale is dominated by safety aspects and a focus on the generation of authentic products. The demand for suitable expression systems is increasing as the emerging systematic genomics result in an increasing number of gene targets for the various industrial branches (for pharmaceuticals, see Chapter 16). So far, the production of approved pharmaceuticals is restricted to *Escherichia coli*, several yeasts, and mammalian cells. In the present book, a variety of expression platforms is described ranging from Gram-negative and Gram-positive prokaryotes, over several yeasts and filamentous fungi to mammalian and plants cells, thus including greatly divergent cell types and organisms. Some of the systems presented are distinguished by an impressive track record as producers of valuable proteins that have already reached the market, while others are newly defined systems that have yet to establish themselves but demonstrate a great potential for industrial applications. All of them have special favorable characteristics, but also limitations and drawbacks – as is the case with all known systems applied to the production of recombinant proteins. As there is clearly no single system that is optimal for all possible proteins, predictions for a successful development can only be made to a certain extent, and as a consequence misjudgments leading to costly time- and resource-consuming failures cannot be excluded. It is therefore advisable to assess several selected organisms or cells in parallel for their capability to produce a particular protein in desired amounts and quality (see also Chapter 13).

The competitive environment of the considered platforms is depicted in Table 1.1. A cursory correlation exists between the complexity of a particular protein and the complexity and capabilities of an expression platform. Single-subunit proteins can easily be produced in bacterial hosts, whereas proteins that require an authentic complex mammalian glycosylation or the presence of several disulfide bonds neces-

Production of Recombinant Proteins. Novel Microbial and Eucaryotic Expression Systems. Edited by Gerd Gellissen

ISBN: 3-527-31036-3

Table 1.1 Some key parameters for the choice of a particular expression system. The column "Expression system" provides the list of the systems described in the various chapters of this book. The column "Classification" provides a rough classification of these organisms. The coloring of the fields indicates the complexity of the respective organism, increasing in the order light gray, medium gray, dark gray. In the following columns, positive and negative aspects are distinguished by the coloring of the fields. Light gray indicates negative, and dark gray positive features. Fields in medium gray indicate an intermediate grading. The column "Development of system" distinguishes between "early stages" and "completely developed". The latter indicates that the full spectrum of methods and elements for genetic manipulations, target gene expression, and handling is available. "Early stages" shall indicate a yet incomplete development. In "Disulfide bonds" and "Glycosylation", two examples of post-translational modification are addressed which may be especially important for heterologous protein production. Prokaryotes have, in general, a strongly limited capability of forming disulfide bonds. If one or more disulfide bonds is necessary for the target protein's activity, a eukaryotic system would be the better choice. If the target protein requires *N*- or *O*-glycosylation for proper function, prokaryotic systems are also disqualified. The production of a glycoprotein for the administration to humans requires special care. So far, only mammalian cells are capable of producing human-compatible glycoproteins. Glycoproteins produced by two methylotrophic yeasts, *Hansenula polymorpha* and *Pichia pastoris*, have been shown not to contain terminal α1,3-linked mannose, which are suspected to be allergenic. For the other yeasts and fungi listed, the particular composition of the glycosylation has yet not been determined, which here is valued as a negative feature. "Secretion" of target protein can be achieved with all systems shown in the list. However, in case of the two Gram-negative bacteria, *Escherichia coli* and *Pseudomonas fluorescens*, "Secretion" means that the product typically accumulates in the periplasm; the complete release requires the degradation of the outer membrane. The following three columns, "Costs of fermentation", "Use of antibiotics", and "Safety costs" refer to a subset of practical aspects for production of a target protein. In general, the "Costs of fermentation" in mammalian cells are much higher than in plant cells, fungi, yeasts, or prokaryotes, due mainly to the costs of the media. However, the use of isopropyl-thiogalactopyranoside (IPTG)-inducible promoters can increase the costs of target protein production in *E. coli* and *P. fluorescens*, as indicated by the medium gray fields. The use of antibiotics in fermentation processes is becoming increasingly undesired. If a therapeutic protein is to be produced in *E. coli* or *Staphylococcus carnosus*, a plasmid/host system should be chosen that allows plasmid maintenance without the use of antibiotics. "Safety costs" refers to the capability of the production system of carrying human pathogenic agents. In this regard, the mammalian-derived cell systems display the highest risks, for example as carriers of retroviruses. "Processes developed" indicates whether processes based on a particular system have already entered the pilot or even the industrial scale, associated with the respective knowledge. "Products on market" indicates which systems have already passed this final barrier.

▶

Expression system	*Classification*	*Development of system*	*Disulfide bonds*	*Glycosylation*	*Secretion*	*Costs of fermentation*	*Use of antibiotics*	*Safety costs*	*Processes developed*	*Products on market*
Mammalian cells	higher eukaryote	completely developed	yes	yes; typically human-like	possible	high	not required	high costs	industrial scale	yes
Plant cells	higher eukaryote	completely developed	yes	yes; terminal fucose	possible; size-restrictions	moderate	not required	low costs	pilot scale	no
Sordaria macrospora	filamentous fungus	earlys stages	yes	yes; exact features yet unknown	possible	low	not required	low costs expected	lab scale	no
Aspergillus sojae	filamentous fungus	completely developed	yes	yes; exact features yet unknown	possible	low	not required	low costs	pilot scale	no
Arxula adeninivorans	dimorphic yeast	early stages	yes	yes; exact features yet unknown	possible	low	not required	low costs expected	lab scale	no
Yarrowia lipolytica	dimorphic yeast	early stages	yes	yes; exact features yet unknown	possible	low	not required	low costs expected	lab scale	no
Pichia pastoris	methylo-trophic yeast	completely developed	yes	yes; no terminal α1, 3 mannose	possible	low	not required	low costs	industrial scale	yes
Hansenula polymorpha	methylo-trophic yeast	completely developed	yes	yes; no terminal α1, 3 mannose	possible	low	not required	low costs	industrial scale	yes
Staphylococcus carnosus	gram-positive bacterium	completely developed	limited	no	possible	low	typically required	low costs	pilot scale	no
Pseudomonas fluorescens	gram-negative bacterium	completely developed	(yes); in the periplasm	no	periplasmic secretion	promoter-dependent low to moderate	not required	low costs	pilot scale	no
Escherichia coli	gram-negative bacterium	completely developed	(yes); in the periplasm	no	periplasmic secretion	promoter-dependent low to moderate	typically required	low costs	industrial scale	yes

sitate a higher eukaryote as host. However, ongoing research and ongoing platform development and improvements might render alternative microbes of lower systematic position suitable to produce such sophisticated compounds. For instance, *E. coli*-based production systems have successfully been applied to a tissue plasminogen activator (t-PA) production process (see Chapter 2); system components are now available for the methylotrophic yeast species *Pichia pastoris* and *Hansenula polymorpha* to synthesize core-glycosylated proteins or those with a "humanized" N-glycosylation pattern (see Chapter 6 on *H. polymorpha*, and Chapter 7 on *P. pastoris*).

Microbial system provide in general easy access to process monitoring and validation as compared to the systems based on higher eukaryotes.

The Gram-negative bacterium *E. coli* was the first organism to be employed for recombinant protein production because of its long tradition as a scientific organism, the ease of genetic manipulations, and the availability of well-established fermentation procedures. However, the limitations in secretion and the lack of glycosylation impose restrictions on general use. Furthermore, recombinant products are often retained as inclusion bodies. Although inclusion bodies sometimes represent a good starting material for purification and downstream procedures, they often contain the recombinant proteins as insoluble, biologically inactive aggregates. This requires in these instances a very costly and sophisticated renaturation of the inactive product. Nevertheless, it still provides the option to produce even complex proteins (as described in Chapter 2), and a range of *E. coli*-derived pharmaceuticals have successfully entered the market.

Pseudomonas fluorescens represents a newly defined system based on an alternative Gram-negative bacterium. Some of the advantageous characteristics of this organism are summarized in Chapter 3, including refraining from antibiotics, improved secretion capabilities, and an improved production of soluble, active target proteins.

Staphylococcus carnosus is a representative of Gram-positive bacteria that are capable of secretion into the culture medium. The platform avoids system-specific limitations frequently encountered with Gram-positive organisms. This includes pronounced proteolytic degradation of products by secreted host-derived proteases, as is the case with commonly applied *Bacillus subtilis* strains. In the case of *S. carnosus*, proteases reside within the cell wall. Potential degradation during cell wall passage can be prevented by using a protective *S. hyicus*-derived lipase leader for export targeting. Additionally, it is possible to secrete lipophilic heterologous proteins that were found to be retained in the insoluble intracellular fraction when using yeasts such as *H. polymorpha*. Another possible application of great potential is the option to tether exported proteins to the surface of the host via C-terminal sorting signal sequences. Recombinant microbes exhibiting such a surface display could be applied to the generation of live vaccines and of biocatalysts (see Chapter 4).

Fungi combine the advantages of a microbial system such as a simple fermentability with the capability of secreting proteins that are modified according to a general eukaryotic scheme. Filamentous fungi such as *Aspergillus* sp. efficiently secrete genuine proteins, but the secretion of recombinant proteins turned out be a difficult task in particular cases. Foreign proteins have to be produced as fusion proteins from which the desired product must be released by subsequent proteolytic processing.

Furthermore, *Aspergillus* usually generate spores that are undesirable in the production of pharmaceuticals. Nevertheless, *Aspergillus* sp. have successfully been used for the production of phytase or for lactoferrin (see Chapter 9). The newly defined *Sordaria macrospora* platform is free of these undesired spores, thereby offering a great potential for the production of recombinant pharmaceuticals (see Chapter 10).

This book also covers a selection of divergent yeast systems. The traditional baker's yeast, *Saccharomyces cerevisiae*, has been used for the production of FDA-approved HBsAg and insulin. Again, severe drawbacks are encountered in the application of this system, and it was therefore excluded from this book: *S. cerevisiae* tends to hyperglycosylate recombinant proteins; N-linked carbohydrate chains are terminated by mannose attached to the chain via a α1,3 bond, which is considered to be allergenic. In contrast, the two methylotrophs harbor N-linked carbohydrate chains with a terminal α1,2-linked mannosyl residue which is not allergenic. Furthermore, the extent of hyperglycosylation is lower as compared to the situation in baker's yeast. Both methylotrophs are established producers of foreign proteins; in particular, *H. polymorpha* is distinguished by a growing track record as production host for industrial and pharmaceutical proteins. Tools have been established in these two species to produce glycoproteins that exhibit a "humanized" glycosylation pattern or that secrete core-glycosylated proteins (see Chapters 6 and 7). More recently, the two dimorphic species *Arxula adeninivorans* and *Yarrowia lipolytica* have been defined as expression platforms. The newly defined systems have yet to demonstrate their potential for industrial processes. Both organisms exhibit a temperature-dependent dimorphism, with hyphae being formed at elevated temperatures. For *A. adeninivorans*, it has been shown that O-glycosylation is restricted to the budding yeast status of the host (see Chapters 5 and 8).

All yeasts – and probably all filamentous fungi – could be addressed in parallel by a wide-range vector for assessment of suitability in a given product development (see Chapter 13).

Mammalian cells [e.g., Chinese hamster ovary cells (CHO) and baby hamster kidney cells (BHK)] are capable of faithfully modifying heterologous compounds according to a mammalian pattern. However, the fermentation procedure is expensive and yields are much lower than those reported for various microbial systems. In addition, mammalian cells are potential targets of infectious viral agents. This forces a vigorous control of all fermentation and purification steps. This situation can be eased to some extent when using hollow-fiber bioreactors, as presented in Chapter 11. To date, the production of industrial compounds is thus restricted to high-price drugs. Nevertheless, very successful pharmaceutical products such as antibodies and their derivatives, or pharmaceuticals such as factor VIII, with its demand for authentic glycosylation, are based on production in mammalian cell cultures.

Plant suspensions cell cultures carry most of the advantages of terrestrial plants, and can be used at present for the production of low or medium amounts of proteins. Benefits include the ability to produce proteins under GMP conditions, the ability to isolate proteins continuously from the culture medium, and the use of sterile conditions. However, further improvements in yield and optimization in downstream processing are required before this platform becomes commercially feasible (see Chapter 12).