Vipin Chandra Kalia · Adesh Kumar Saini
Editors

Metabolic Engineering for Bioactive Compounds

Strategies and Processes

Springer

Vipin Chandra Kalia • Adesh Kumar Saini
Editors

Metabolic Engineering for Bioactive Compounds

Strategies and Processes

Springer

Editors
Vipin Chandra Kalia
Microbial Biotechnology and Genomics
CSIR - Institute of Genomics and Integrative Biology (IGIB)
New Delhi
India

Adesh Kumar Saini
Faculty of Applied Sciences and Biotechnology
Shoolini University of Biotechnology and Management Sciences
Solan, Himachal Pradesh
India

ISBN 978-981-10-5510-2 ISBN 978-981-10-5511-9 (eBook)
DOI 10.1007/978-981-10-5511-9

Library of Congress Control Number: 2017951913

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

Editors' Biography

Vipin Chandra Kalia is an emeritus scientist of the CSIR Institute of Genomics and Integrative Biology, Delhi. He holds an M.Sc. and Ph.D. in Genetics, both from the Indian Agricultural Research Institute, New Delhi. Further, he is a fellow of the National Academy of Sciences, National Academy of Agricultural Sciences, and Association of Microbiologists of India.

His main areas of research are microbial biodiversity, bioenergy, biopolymers, quorum sensing and quenching, and drug discovery. He has published more than 100 papers in scientific journals including *Nature Biotechnology*, *Biotechnology Advances*, and *Bioresource Technology*.

He has authored 14 book chapters. He has edited seven books and is currently the editor in chief of the *Indian Journal of Microbiology* (since 2013) and editor of several journals, including the *Journal of Microbiology and Biotechnology* (Korea) and *Dataset Papers in Microbiology*. He is a member of numerous professional associations, including the Society of Biological Chemists of India and the Biotech Research Society of India (BRSI). He can be contacted at vckalia@igib.res.in and vc_kalia@yahoo.co.in.

Adesh Kumar Saini is a professor and the director of the Center of Research on Himalayan Sustainability and Development, Shoolini University, India. He received his Ph.D. in Microbiology from the CSIR Institute of Genomics and Integrative Biology and completed his postdoctoral research at the National Institutes of Health. His main areas of research are microbial biodiversity, sustainable development of the Himalayan ecosystem, and yeast cell biology. He has published more than 25 papers in prominent scientific journals, including *The EMBO Journal* and the *Journal of Biological Chemistry*. His research projects aim at plant-growth-promoting rhizobacteria (PGPR), postfire management in the pine forests, and revival of forest productivity. He can be contacted at sainiade@gmail.com and adeshsaini@shooliniuniversity.com.

Contributors

Anshu Alok Department of Biotechnology (DBT), National Agri-Food Biotechnology Institute, Mohali, Punjab, India

Meenakshi Balhara Centre for Biotechnology Maharshi Dayanand University, Rohtak 124001, Haryana, India

Saurabh Bansal Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Gopal Singh Bisht Department of Pharmacy, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Renu Chaudhary Centre for Biotechnology, Maharshi Dayanand University, Rohtak, Haryana, India

Anil Kumar Chhillar Centre for Biotechnology, MD University, Rohtak, Haryana, India

Anil A. Chuturgoon Discipline of Medical Biochemistry and Chemical Pathology, School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

Andrés Gárriz Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Naseem A. Gaur Yeast Biofuel Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi, Delhi, India

Ritu Ghildiyal Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

María Elisa Gonzalez Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/ Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Kinam Gupta Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Vipin Chandra Kalia Microbial Biotechnology and Genomics, CSIR—Institute of Genomics and Integrative Biology (IGIB), New Delhi, India

Savita Khatri Centre for Biotechnology, Maharshi Dayanand University, Rohtak, Haryana, India

Jitesh Kumar Department of Biotechnology (DBT), National Agri-Food Biotechnology Institute, Mohali, Punjab, India

Manish Kumar Centre for Biotechnology, MD University, Rohtak, Haryana, India

Manoj Kumar Amity Institute of Microbial Technology, Amity University, Noida, Uttar Pradesh, India

Reena Kumari Animal Biotechnology Laboratory, Faculty of Applied Sciences and Biotechnology, Shoolini University, Solan, Himachal Pradesh, India

Bisma Malik Department of Bioresources, University of Kashmir, Srinagar, Jammu and Kashmir, India

María Marina Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Deepak Pathania Department of Environmental Sciences, Central University of Jammu, Jammu, Jammu and Kashmir, India

Neetu Phougat Centre for Biotechnology, MD University, Rohtak, Haryana, India

Fernando Luis Pieckenstain Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/ Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Tanveer Bilal Pirzadah Department of Bioresources, University of Kashmir, Srinagar, Jammu and Kashmir, India

Poonam Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Ashu M. Poswal Faculty of Applied Sciences and Biotechnology, Shoolini University, Solan, Himachal Pradesh, India

Reiaz Ul Rehman Department of Bioresources, University of Kashmir, Srinagar, Jammu and Kashmir, India

Fernando Matias Romero Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/ Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Franco Ruben Rossi Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/ Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Adesh K. Saini Faculty of Applied Sciences and Biotechnology, Shoolini University of Biotechnology and Management Science, Solan, Himachal Pradesh, India

Reena V. Saini Animal Biotechnology Laboratory, Faculty of Applied Sciences and Biotechnology, Shoolini University, Solan, Himachal Pradesh, India

Jata Shankar Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Arush Sharma School of Chemistry, Shoolini University, Solan, Himachal Pradesh, India

Deepika Sharma Department of Pharmacy, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Shivam Sharma Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Rahul Shrivastava Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Hemant Sood Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Swadeep Sood School of Chemistry, Shoolini University, Solan, Himachal Pradesh, India

Ankit Srivastava Kusuma School of Biological Sciences, Indian Institute of Technology, New Delhi, Delhi, India

Raman Thakur Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan, Himachal Pradesh, India

Charlette Tiloke Discipline of Medical Biochemistry and Chemical Pathology, School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

Department of Chemistry, Faculty of Applied Sciences, Durban University of Technology, Durban, South Africa

Subodh Verma National Institute of Plant Genome Research, New Delhi, Delhi, India

Paulina Vignatti Laboratorio de Interacciones Planta-microorganismo, Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús/ Universidad Nacional de General San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas (IIB-INTECH/UNSAM-CONICET), Chascomús, Buenos Aires, Argentina

Farnaz Yusuf Yeast Biofuel Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi, Delhi, India

Developments and Diversity of Proteins and Enzymes

2

Ankit Srivastava, Saurabh Bansal, and Jata Shankar

Abstract

Proteins form an integral part of a large consortium of industrially and therapeutically important products. Thus, efficient strategies for producing high-quality proteins and enzymes at industrial scale have gained importance in the past few decades. However, the heterogeneous production of recombinant proteins at industrial setups requires remotion and optimization of certain drawbacks such as protein aggregation, contamination, unstable cultures and less productive fermentation process. Therefore, the ultimate aim is to develop bioprocess that could be utilized for efficient production of important proteins. This has led to the development of structural biology, protein engineering and bioinformatics tools for identifying and rectifying these issues at molecular level. In the postgenomic era, a huge variety of protein families are yet to be functionally characterized. In this context, the discovery of new proteins with varied functionalities can only be achieved using modern-day methodologies that incorporate sequences, structure and activity-based approaches. Later, the positive outcomes could be easily directed towards the development of stable recombinant culture and productive fermentation process alongside a cost-effective downstream process. This chapter discusses the strides in the protein/enzyme production and their biotechnological perspectives along with the limitations in scale-up. Besides, several methodologies for selecting efficient protein molecules such as directed evolution approaches as well as for studying their physicochemical characteris-

A. Srivastava
Kusuma School of Biological Sciences, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India

S. Bansal (✉) • J. Shankar
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan 173234, India
e-mail: saurab.bansal02@gmail.com

V.C. Kalia, A.K. Saini (eds.), *Metabolic Engineering for Bioactive Compounds*,
DOI 10.1007/978-981-10-5511-9_2

tics such as single-molecule techniques are also described. In addition, this chapter also delineates production of therapeutic proteins and proteases and applications of chaperones in industrial production process.

Keywords

Therapeutics • Bioprocess • Directed evolution • Protein engineering • Interactions Single-molecule biophysics • Fluorescence • Atomic force microscopy • Biochemical diversity • Protease • Down-stream processing • Chaperones • Protein folding

2.1 Protein Science: Journey from Cell to Industry

The past two decades had been crucial for the development of industrial production process of recombinant proteins. Several new strategies have been developed and implemented for producing large amounts of high-quality proteins of varied commercial applications. A huge number of new proteins from a range of origins such as virus, bacteria, fungus, plants and higher animals have been produced using genetically modified organisms (GMO) (Ahmad et al. 2012; Buiatti et al. 2013). The postgenomic era has opened new vistas in identifying alternate sources of scarce proteins which were earlier found difficult to produce at small or industrial scales. These proteins/enzymes are now readily produced in forms resembling closely to the actual native proteins. This has also cut down the amount of protein and protein-based products that were earlier obtained directly from the human or animal tissues. During the past decade, the US Food and Drug Administration (FDA) has been approving 15 new recombinant proteins on an average each year (Mullard 2014). The protein production at industrial scale differs from traditional small molecules that are chemically synthesized. In this case, recombinant DNA technology plays the pivoting role along with protein expression systems directed in various living systems. Basically, successful protein/enzyme production process depends heavily upon the following five factors, viz. (a) production cost, (b) control on actual product processing such as native post-translational modifications, (c) processing time for producing protein from the gene of interest, (d) host-based nuisances in protein production and (e) control of protein aggregation (Rai and Padh 2001; Gupta and Shukla 2015).

The postgenomic boom in genetic information has led to an enormous increase in a number of innovative biological products and biosimilars that mostly comprise proteins. This has fuelled the demand of customized production cell lines with high level of productivity. Nevertheless, heterologous protein production at industrial scale still requires much optimization and differs significantly case by case. Based on widely characterized genetic architecture, *E. coli* has been a widely employed host in protein products that do not require post-translational modifications (Chen 2012). Besides safe genetic manipulations, the simplified purification strategies for producing recombinant proteins using *E. coli* at industrial scales also make them a preferred host. On the other hand, several protein products have been industrially produced using yeast as an alternative host (Celik and Calık 2012). This host is typically free from pyrogens and viral and

oncogenic DNA and thus is a preferred for production of pharmaceutical proteins (Mattanovich et al. 2011). Apart from this, baculoviruses have evolved as popular systems for recombinant protein overproduction in eukaryotic cells (Fornwald et al. 2016). Owing to the similar cellular milieu, these insect cell lines promote protein modifications and processing similar to higher eukaryotic cells and thus act as preferred hosts for producing qualitatively pure proteins (Radner et al. 2012). These are generally employed when the product specificity is essential and aids in its clinical efficacy. Collectively, the design of heterologous protein production process depends heavily on the level of purification, stabilization and formulation required for a particular protein product.

The other major drawback in industrial protein production is attributed to protein aggregation and precipitation (Vázquez-Rey and Lang 2011). Protein folding is a complex phenomenon that incorporates two major components, i.e. foldases, that accelerate folding and chaperones that prevent formation of insoluble protein aggregates. In an industrial bioprocess, due to the presence of numerous cell stress factors such as heat shock, nutrient depletion or other environmental contaminants, the protein products tend to aggregate (den Engelsman et al. 2011). These aggregates due to their toxic nature further lead to loss in biomass and hence significantly affect the efficacy of the production process. Thus, recovery and renaturation steps are added to obtain optimized amounts of proteins/enzymes in an economically feasible fashion. Thus, several molecular techniques such as utilization of protease-deficient strains as well as bioprocess strategies such as high-temperature culturing and engineering of fusion peptides/proteins that promote protein solubilization have been widely used (Clark 2001; Rajan et al. 2011; Levy-Sakin et al. 2014). Incorporation of all these schemes into a unified bioprocess scale-up will lead to achieve the goal of obtaining high-quality proteins/enzymes in excellent quantities.

2.2 Protein Families: From Biotechnology Perspective

The biotechnological advances in recent times have been found to heavily depend upon new sequencing technologies and advances in proteomics that have enabled large-scale identification and functional characterization of new proteins and enzymes (Ideker et al. 2001; van Beilen and Li 2002; Hartmann et al. 2014). Nevertheless, the process of identification is often complicated owing to the fact that sequence similarities do not warrant similar functional characteristics. Since, a small modification or mutation in a protein could severely affect its structure and native functionality (Wang and Moult 2001). Besides, several microbes with identified functions may not essentially harbour an expected gene in their genomes. Also, there are many protein families that are yet to be discovered based on sequence-based or activity-based approaches. The newly sequenced genomes have created ample sequence information about umpteen of proteins/enzymes spread across the taxa (Loman et al. 2012). Thus, it is essential to annotate these genes and understand how these proteins function in the living cells. This would essentially result in novel protein candidates that could be used in the biotechnological perspective. Since the experimental verification requires time as well as large amount of technical infrastructure, recently several computational algorithms have been developed to partially predict the functions of unknown proteins. Although highly

error prone, the bioinformatics approaches essentially cut down the large list and help in creating a precise catalogue of gene products (Weckwerth 2011). Thus, the discovery of new enzymes and proteins helps in cataloguing functional homologues that could serve as alternatives to the proteins that are industrially and therapeutically important. This essentially is an important prerequisite for fuelling research in protein science for identifying new therapeutics as well as commercially important proteins/enzymes.

A large number of environmental factors influence the survival of microorganisms with fitting genetic machinery that helps them sustain. Extreme environmental stress directs the evolution of thermostable, psychrophilic and halophilic proteins as well as proteins with specific functions such as degradation of xenobiotics and other chemical substances (Elleke et al. 2013). In the recent past, several databases and automatic annotation tools have been developed that include IMG-M, MG-RAST, COG/KOG, TIGRFAM, CDD and PROSITE (Kunin et al. 2008). The new protein families' characterization and the novel functionalities identification of the already existing protein families are the major outcome of an integrated pipeline consisting of biochemical, structural, genomic and meta-omics data (Thomas et al. 2012). This pipeline essentially helps in characterizing huge number of newly identified scaffolds that await functional assignment. In the coming years, these strategies would help in identifying novel proteins and enzymes with varied functions that could be developed into a wide variety of biotechnological products such as therapeutic molecules, drug delivery scaffolds, enzymes and various other commercial products.

2.3 Protein Production: Scaling Up and Limitations

Since proteins are important as therapeutics, as catalysts and as diagnostics tools, their demand has continuously increased (Tripathi et al. 2009). Therefore, the ultimate goals of the development of bioprocess and its scale-up are:

1. The production of desired proteins at minimal cost
2. Maximize the volumetric productivity, i.e. the maximum amount of protein obtained in given volume in the least period of time

Thus, nowadays, most of the proteins are the recombinant proteins obtained through heterogeneous production (Xia et al. 2015). The heterogeneous production of recombinant proteins requires a genetically and biochemically stable recombinant culture with higher yield, a high productive fermentation process and cost-effective downstream process. Various strategic choices may be used to the process for enhancing the protein production while designing the purification protocols and strategies, and these can be listed as follows:

1. Selection of right expression host for the protein(s): bacteria, yeast, insect cells or human cells.
2. Selection of suitable expression vector.
3. Selection of suitable strain(s) if bacterial expression is used.
4. Whether the expression of full-length protein is required or a fragment thereof.

5. Whether the tagging of protein is required. If yes, which one is the best?
6. Identification of a good purification strategy and its common pitfalls.

Since every protein is in itself different, unfortunately, the answer to the each of these questions a priori cannot be fitted with every protein (Structural Genomics Consortium et al. 2008). Therefore, the strategies and the purification protocols need to be designed individually for each protein while keeping in mind of its intended use. The above given questions have already been discussed earlier in the book. After fixing all these queries, the next step is the scaling up of the process. Thus, the development of successful scaling-up process is the key factor for the commercial success of any proteins.

The scale-up and optimization of the protein production via biochemical activities of microorganisms, including bacteria, yeasts, algae and moulds, and of animal and plant cell systems require sophisticated engineering skills. To achieve the protein production at commercial level needs the increase in size of fermenter without affecting the overall production efficiency of the process. For this purpose, various fermentation parameters need to be considered which are directly linked with the biochemical activities of the culture and so the protein production as the protein is directly linked with the biochemical activities of the cells. Majorly mixing, mass transfer and heat transfer play the critical role during scale-up of the process (Schmidt 2005). Theoretically the following criteria are assumed suitable bases for the scaling up of the process (Ju and Chase 1992):

1. Constant fermenter geometry (constant height to diameter ratio; *H*/*D*)

   While selecting constant H/D as a basis for the scale-up, the surface to volume ratio declines dramatically during scale-up. It will decrease the relative contribution of surface aeration to oxygen supply and dissolved carbon dioxide removal as compared to sparging contribution. This factor can be critical when the shear-sensitive cultures like animal cells are being used for the protein production which leads to significant decrease in the volumetric productivity. In contrast, this factor will not affect the protein production from the microbial culture as in this case the surface aeration is unimportant (Shuler and Kargi 2002).
2. Constant power per liquid volume

   To minimize the cost of final protein product and to earn the maximum profit, the protein production at commercial level must be cost-effective which is directly linked with the power consumption during aeration and mixing of the system. Therefore, while scaling up, constant power per liquid volume can be an important factor. For geometrically similar, fully baffled vessels with turbulent conditions, it may be noted that if scale-up should be based on maintaining a constant power input per unit volume considering no gassing in the system, one may have the following:

$$N_1 = N_2 \left(D_2 / D_1 \right)^{2/3}$$

where N = impeller speed, D = tank diameter and subscripts 1 and 2 represent lab scale and scaled-up fermenter.

For the gassing system,

$$N_1 = N_2 \left(D_2 / D_1\right)\left(Q_1 / Q_2\right)^{1/14}$$

where Q = volumetric gas flow rate.

3. Constant volumetric mass transfer coefficient (K_La)

   Volumetric mass transfer coefficient is an important criterion for an aerobic fermentation. Most of the protein production is carried out under the aerobic condition. During the scale-up of a process for the aerobic fermentation, K_La needs to be increased in the way which it can fulfil the oxygen demand of the biomass during protein production. In taking constant K_La as basis of scale-up criteria, the following equation can be used during the scaling up of fermenter for the protein production:

$$N_1 = N_2 \left(D_2 / D_1\right)^{13/21} \left(Q_2 / Q_1\right)^{5/42}$$

4. Constant impeller tip speed

   The impeller tip speed is important not only for the well mixing of the reactor content but also for the culture protection from shear damage. If the animal cell culture or the bacterial cells are the source for the protein production and as these cells are sensitive to the shear damage, this factor needs to be optimized carefully while scaling up the process; otherwise, it may lead to the less protein production due to the loss of active biomass. Therefore, the equal impeller tip speed in both the laboratory and full-scale plant reactor is one of the key scale-up criteria. The typical impeller tip speed ranges from 5 to 7 m/s:

$$N_1 = N_2 \left(D_{i2} / D_{i1}\right)$$

where D_i = impeller diameter.

5. Constant mixing quality (constant Reynolds number)

   Constant Reynolds number in other words indicates the constant mixing quality. The constant mixing quality for the optimum protein production is important as it affects both the mass and heat transfer in the system which may lead to be a limiting factor for the biomass production and so protein production. Therefore, constant Reynolds number can also be used as the basis criteria for the scale-up:

$$t_{m2} = t_{m1} \left(D_2 / D_1\right)^{11/18}$$

where t_m = mixing time.

6. Constant mixing rate number

Minimum mixing time is desirable for creating the homogenous environment in the reactor. Mixing time is often used as scale-up criteria. Mixing time (t_m) can be defined as the minimum time required for the homogeneous distribution of a small volume of pulsating masses in the bulk of the liquid. It is important to adjust the proper mixing conditions in the large vessels during the scale-up. However, the use of mixing time as a scale-up criteria basis has its own limitations as to maintain the t_m constant in both the vessels increases the power expenditure significantly which is not suggestible as it increases the production cost significantly. Thus, as a whole, it can be understood now that the major limitations of scale-up are the mass, heat and momentum transfer deficiencies with the increase in size of the bioreactors which affect the overall production of the protein (Palomares and Ramírez 2003). Mass transfer becomes more important when the nutrient or chemical inducers like IPTG are to be added while the fermentation is on run. With the increasing size of reactor, mixing time increases significantly (Junker 2004). For example, mixing time in large bioreactors containing animal or plant cell culture (10,000 l) can be in the order of 10^3 s. Thus, approximately 17 min would be required for the homogenous distribution of nutrient or inducer in the reactor under the specified conditions. However, it can be solved by adding these chemical inducers or nutrients though multiple ports to the reactors (Ozturk 1996).

Other important mass transfer limitation is related to the oxygen supply to the cells in aerobic culture system due to low solubility of oxygen in water (Shuler and Kargi 2002). The problem gets aggravated with the increase of cell densities due to higher oxygen demand for supporting the cell growth as compared to the oxygen supply. The effective mixing in reactor to generate homogenous environment is necessary to enhance the volumetric productivity of the protein; otherwise, it may lead to the alcoholic or acid fermentation in bacteria, yeast and animal cell cultures (Palomares and Ramírez 1996; O'Beirne and Hamer 2000). Therefore, during scaling up of the process, the bioreactor should be designed to have a higher K_La to increase the oxygen transfer rate (OTR) as much as possible. In general, the dissolved oxygen tension (DOT) should be higher than 20% (with respect to air saturation) to support the growth of the culture. The limitation of oxygen transfer in bulk liquid to the cell may also be due to diffusion through additional resistances, such as cell aggregates or pellets and immobilized cells. In such cases, a 50% DOT in the bulk liquid may be required to sustain the growth of aggregated or immobilized cells in the reactor (Yegneswaran et al. 1991).

As a whole, the primary limitations of the scale-up are gas supply and heat removal. For geometrically similar tanks, shear, mixing time and K_La are impossible to maintain identically simultaneously in both the large and small tanks. Thus, scale-up is difficult as various gradients such as DO gradient, substrate gradient, temperature gradient, pH gradient and CO_2 gradient do occur in the large reactor which led to the induction of multiple physiological responses by the cells. Thus, such oscillating environment in large reactors leads to a heterogeneous cell growth. Consequently, biomass, productivity and yield get lowered

significantly, and the by-product formation gets enhanced as compared to small-scale reactors (Bylund et al. 1998; Enfors et al. 2001; Käß et al. 2014). Therefore, a successful conversion of lab scale process to commercial scale requires a lot of time and efforts which is the major limitation for the rapid development of a successful scaling-up process (Neubauer et al. 2013).

2.4 Directed Evolution Approaches

Native proteins many a times are required to be modified and altered to an extent that they meet the industrial demands. Directed evolution has been long proven as an effective approach for enhancing the properties of proteins for commercial and therapeutic applications (Cobb et al. 2012; Goldsmith and Tawfik 2012; Porter et al. 2016). This technique relies on evolution of proteins in laboratories through the manifestation of diverse environmental strains. This is followed by screening and identification of protein variants with desired properties (Kumar and Singh 2013). During the course of evolution, iterative mutations and natural selection have helped organisms to survive in the changing environments. To direct useful phenotypes by providing artificial environmental constraints forms the base of directed evolution approaches. This artificial selection process has been utilized since long for selective breeding of crops as well as domestication of animals. Thus, directed evolution has been established as a highly effective and broadly applicable methodology for obtaining modified gene products. This technique basically mimics the natural evolution process and has been hence termed as 'directed evolution' or 'in vitro evolution' (Zhao and Zha 2004) (Table 2.1).

Earlier, the traditional genetic screens utilized physicochemical agents that randomly modified genomic DNA. This methodology was called random mutagenesis that relied upon agents that mainly included several chemical compounds such as ethyl methanosulphonate (EMS), nitrous acids, 2-aminopurines and even ultraviolet (UV) radiations (Kaur and Sharma 2006). These artificial mutagenesis agents actually augment the rate of mutations during the DNA replication process that result in modified gene products. Later, due to relatively lower mutation rates and lack of controls, in vivo random mutagenesis strategies were employed which are still in infancy (Ravikumar et al. 2014). Further, several workers reported the use of error-prone PCR (epPCR) that utilizes DNA polymerases with mutational bias resulting in a large mutational scope (Leung et al. 1989; Camps et al. 2003). Another important methodology that was developed later included DNA shuffling (Stemmer 1994). This was the first homologous recombination method that involves digestion of particular gene by DNases followed by the reassembly of fragments by PCR. Although the fragments reassemble based on sequence homology, the entire gene also incorporates recombinations during the crossover events (Coco et al. 2001). Owing to high rate of point and deletion mutation, a library of mutants could be generated for screening new mutants. Later, a modification to this technique was developed that utilizes recombination of a family of related genes from various species (Crameri et al. 1998). This was termed as family shuffling and involved

Table 2.1 Comparison of different direct evolution approaches

Approach	Methods	Advantages	Disadvantages
Chemical mutagenesis	Ethyl methanesulphonate (EMS), nitrous acid, UV radiation, etc.	Mutations are dose dependent	Chemical hazard, low and uneven mutations
Mutator strains	Phage-assisted continuous evolution (PACE), XL1-red *E. coli*, etc.	Comparatively simple	Low and uneven mutations
Error-prone PCR (ep-PCR)	Taq polymerase with varying dNTPs and divalent ion (Mg^{2+}, Mn^{2+}) concentrations, etc.	High and even mutation rate	Non-even sampling of codons
Homologous recombinations	DNA shuffling, family shuffling, StEP, RACHITT, NExT, etc.	Ease of identification of useful or futile mutations Sequence reshuffling for attaining functional diversity in protein families	Overdependence on sequence homology
Nonhomologous recombinations	ITCHY, SHIPREC, NRR, overlap extension PCR, etc.	Ease of shuffling poorly related sequences.	Poor general applicability

StEP Staggered extension process, *RACHITT* random chimeragenesis on transient templates, *NExT* nucleotide exchange and excision technology, *ITCHY* incremental truncation for the creation of HYbrid enzymes, *SHIPREC* sequence homology-independent protein recombination, *NRR* non-homologous random recombination

naturally occurring homologous genes. The technique was found effective in producing improved enzymes as well as structurally stable proteins using a single recombination-selection cycle. More recently, focussed mutagenesis approaches have been developed that utilize synthetic DNA oligonucleotides which contain one or more degenerate codons for targeted sites (Acevedo-Rocha and Reetz 2014). These mutation-inducing oligonucleotides are incorporated into a gene library using restriction digestion or gene assembly protocol. This results in simultaneous saturation mutagenesis that has higher probabilities of recombination and an increased ability to produce a large library of mutants that could be screened of desired property.

With further advent of technology, random chimeragenesis on transient templates (RACHITT) was developed which was an improved version of the DNA shuffling technique (Coco et al. 2001). In this case, fragmented DNA anneals directly to a transient polynucleotide scaffold which is further digested. This helps in increasing number of fragment recombinations as well as crossover frequencies. Further modification to this technique in the form of nucleotide exchange and excision technology (NExT) has been proposed that incorporates treatment with uracil-DNA-glycosylase (UDG) and piperidine or apurinic/apyrimidinic lyases leading to higher frequency of recombinations and larger library sizes (Müller et al. 2005). Further, several nonhomologous recombination methods have also been devised that help in shuffling genes with significantly lower sequence identity. The major

technique is the incremental truncation for the creation of hybrid enzymes (ITCHY). Here, the domains from two parent enzymes are randomly fused to generate chimaera library that is screened later for novel activities. The technique involves exonuclease III for the controlled digestion of DNA to generate truncated fragments of parent genes in maximum numbers which further undergoes blunt-end ligation to form chimaeras (Ostermeier et al. 1999).

Several screening procedures for directed evolution approaches have been also developed alongside. Spatial separation screening involves expression in model unicellular organisms such as *E. coli* or yeast followed by screening as colonies against environmental constraints. Besides this, spatial separation technique has also been recently utilized for screening individual mutants. Several workers have utilized fluorescent readouts, high-performance liquid chromatography (HPLC), nuclear magnetic resonance (NMR), gas chromatography as well as mass spectroscopy for directly monitoring substrate utilization or product formation in colonies (Cobb et al. 2012). Apart from these, earlier fluorescence-activated cell sorting (FACS) had also been used for screening mutant colonies (Fulwyler 1965). The technique relies on a non-diffusing fluorescent reporter (fluorescent-labelled antibodies or tags) for the identification and isolation of cells containing desired gene variants from heterogeneous cell mixture. The current directed evolution protocols have been effectively utilized for enhancing solubility, substrate affinity, catalytic efficiency, thermostability and activity for multiple substrates in a large variety of proteins.

2.5 Single-Molecule Studies of Proteins

Force represents a crucial functional and structural parameter that defines the role of biological macromolecules in their native environment. This has made nanomechanical single-molecule experiments as important tools for probing variety of biomolecular events such as protein folding, protein-ligand interactions, DNA-protein interactions, molecular motors and enzyme catalysis (Mehta et al. 1999; Deniz et al. 2008). These experimental tools amalgamate material sciences with molecular sciences for precise measurement of miniscule regulatory forces that determine function in the life forms. The outcome gives important insights into the intra- and intermolecular interactions that influence biological macromolecules. Looking at the historical perspective, earliest reported single-molecule experiments were the patch clamp-based conductance measurement of ion channels (Sakmann and Neher 1984). However, this technique has its limitations in the choice of systems that could be studied. In due course, spectroscopy provided a more versatile and convenient means to study single molecules (Joo et al. 2008). The commercial availability of sensitive optical detectors made it easier to monitor molecules that emit photons upon excitation. Besides, this also helped in precise observation of photon emission events with high sensitivity and time resolution. Thus, single-molecule fluorescence-based studies became more relevant with the advent of better imaging and biomolecule labelling techniques (Tian et al. 2017; Chen and Ting 2005; Peterman et al. 2004). Later, owing to its functional operation at biomolecular length

Table 2.2 Comparison of single-molecule techniques used for studying biomolecules

Method	Principle technique	Force (nN)	Resolution (nm)	Applications
Single-molecule FRET	Fluorescence	None	2–10	Protein folding and interactions
Magnetic tweezers	Electromagnet and magnetic bead	0.001–100	1–10	Stretching and deformations in DNA and proteins
Optical tweezers	Infrared laser and microbead	0.1–150	0.1–5	Molecular motors, actin assembly
Microneedles	Optical fibre	1–1000	10	DNA and protein deformations
Atomic force microscopy (AFM)	Cantilever and piezoelectric surface	5–1000	0.1–1	Protein folding, interactions, DNA super structures, DNA unwinding
Biomembrane force probes (BFP)	Lipid vesicle and suction pipette	0.01–1000	10	Protein-ligand and protein-lipid interactions

scales, atomic force microscopy was also widely employed in single-molecule-based studies. Currently, several techniques exist that can take into account a wide range of biomolecular forces. These include optical tweezers, glass microneedles, magnetic beads and biomembrane force probes (BFP) (Table 2.2). Thus, single-molecule study has gradually emerged as an important tool for investigating structure-function correlation of biomolecules.

2.5.1 Single-Molecule FRET

Current understanding of protein folding phenomenon has taken an interesting turn with the advent of single-molecule fluorescence techniques. Single-molecule FRET (Förster resonance energy transfer) represents a very precise and controlled methodology to understand the folding of individual protein molecules (Haran 2003). The relaxation rates observed during the folding kinetics of proteins give only the theoretical steps defining the folding states of a protein (Michalet et al. 2006). Considering a protein exhibiting simple two-state folding behaviour, it can be assumed that the measured spectroscopic property (during unfolding/refolding) is a linear combination of averaged values that demarcate conformational states populated by folded and unfolded states at different time points. On the other hand, theoretical studies provide direct information about the behaviour of single protein molecule under the same experimental condition (Snow et al. 2005). Thus, the single-molecule FRET experiments are aimed at direct visualization of folding events that directly correspond to an individual protein moiety. The methodology involves the attainment of trajectories of FRET efficiency or the distance versus time plot during the time scale of transition between unfolded and folded states or vice versa (Lee et al. 2005).

Unlike other spectroscopic techniques, FRET provides quantitative details of the variation in intermolecular distances. Thus, FRET is sometimes called the 'spectroscopic ruler' in the areas of biophysical studies on biomolecules (Piston and Kremers 2007). Nevertheless, doubts on preciseness of this distance calculation as well as efficiency in multiple experiments remained. This incertitude was dealt in detail by the works utilizing polyproline as spacer between donor and acceptor chromophores. It was observed that the average FRET efficiency at shorter distances was slightly lower than theoretically calculated by Forster theory and way much higher for the longer distances (Watkins et al. 2006). The plausible explanation to this divergence could be the negligible information about orientational averaging during donor lifetime as well as the requirement of interchromophore distance longer than the sizes of chromophores. This early evidences paved way for the development of more accurate single-molecule FRET methodologies. Recently, more precise calculations indicated that molecular simulations could be used in alignment with the FRET-based measurements (Reif and Oostenbrink 2014; Best et al. 2015).

A classical single-molecule protein folding experiment monitored by FRET identifies photon bursts from freely diffusing protein molecules under increasing denaturant concentration. The major advantage of this simple experiment is that it is free from any artefacts that may arise from surface interactions. Each subpopulation of unfolded, partially folded and folded protein moieties slowly interconvert, and the FRET efficiency distribution displays peaks corresponding to the mean efficiency of each subpopulation (Haran 2003; Forman and Clarke 2007). Basically, the peak numbers in the FRET efficiency distribution precisely show the number of thermodynamic states during this simple protein refolding/unfolding experiment. In most efficiency distribution experiments consisting of slowly refolding and unfolding rates, only two major peaks are obtained at an intermediate denaturant concentration. These two peaks represent the unfolded and folded subpopulations of the protein. Similarly, several protein systems have been studied that include cold shock protein (CspTm) (Nettels et al. 2007), Acyl-CoA-binding protein (ACBP) (Laurence et al. 2005), Chymotrypsin inhibitor 2 (CI2) (Deniz et al. 2000), RNase HI (Kuzmenkina et al. 2006) and immunity protein Im9 (Tezuka-Kawakami et al. 2006). Further, workers have also utilized other applications such as microfluidic devices in association with single-molecule FRET measurements to elucidate the protein folding phenomenon more precisely (Lipman et al. 2003).

2.5.2 AFM-Based Force Spectroscopy

Biomolecular interactions are majorly regulated by multiple weak but specific forces that define the affinity of interacting partners. Single-molecule studies have identified topological features and mechanical characteristics of these forces and have thus provided mechanistic insights into numerous biological interactions. With the advent of scanning probe microscopy (SPM) and later the atomic force microscopy (AFM), force spectroscopy emerged as a new tool for studying biomolecular systems (Liu and Wang 2011). The AFM utilizes a sharp tip of a flexible cantilever

that probes the mechanical characteristics as well as numerous other parameters of a sample (Carvalho and Santos 2012). Most current day AFM machines rely upon computer-controlled piezoceramic stage that holds the sample and the cantilever that moves over the sample. The deflection force acting between the sample surface and the cantilever is utilized to precisely calculate topological and mechanical features of the sample. This deflection is generated due to lateral (XY) and vertical (Z) displacements of the sample with respect to probe (Dufrene 2002). The sensitivity and resolution of the calculated force depend on the spring constant of cantilever and the laser power that monitors the movement of cantilever. There are two general modes of cantilever motion, 'tapping' mode and the 'contact' mode. In 'tapping' mode, the cantilever is oscillated above the sample surface near its resonance frequency, whereas the 'contact' mode majorly involves constant approach-retraction cycle of cantilever close to sample surface (Garcia and Perezb 2002).

The cantilever motion is utilized to calculate the force, and thus a typical force versus displacement (FD) curve is generated. The variation of force in an FD curve describes the movement of probe towards the sample surface. The movement of probe close or away from the surface is equated with the elastic force of the surface that defines the interaction of the probe and the sample (Anczykowski et al. 1999; Butt et al. 2005).

This interaction is defined on the basis of Hooke's law:

$$F = -k\mathrm{d}$$

where F is the force acting on the cantilever, k is the spring constant of the cantilever and d is deflection.

The cantilever force constants are calibration dependent and vary with different materials used to prepare the probes. Basically, the probe/sample separation affects the different forces that act on the cantilever. Several complex phenomena such as receptor-ligand, protein-protein, DNA-ligand and DNA-protein interactions have been analysed using single-molecule force spectroscopy measurements (Churnside et al. 2012). The biotin/streptavidin interaction is a classical ligand/receptor interaction that has been studied in detail using AFM-based force spectroscopy. The unusually high affinity of the pair as well as the availability of structural information of the interaction has aided in precise calculation of forces acting between them (Grubmuller et al. 1996; Yuan et al. 2000). Several groups have studied the variation in these forces by using mutant protein as well as different analogues of biotin. Besides, this intermolecular force measurements involving protein/protein interactions have been greatly studied using numerous antigen/antibody interactions (Carvalho and Santos 2012). In this case, owing to the diversity in protein pairs, there were tremendous variations in the force measurements. The resulting information provided ample information about the acting forces in different protein scaffolds and molecular pairs and their correlation with their physiological functions.

Another complex phenomenon studied using single-molecule force spectroscopy is the protein folding/unfolding. Several research groups have utilized AFM to study the stretching of single- or multi-domain proteins (Fisher et al. 1999; Besta et al. 2003). In these cases, the FD curves were correlated with the enthalpic and

entropic parameters associated with protein folding. In most cases, the proteins generate initial entropic stretching that further leads to domain-wise unfolding generating continuous cycles of deflections. This cycle of deflections is followed by entropic stretching everytime a new domain is unfolded. The compilation of entire entropic and enthalpic cycles is used to define folding/unfolding pathway of the protein.

2.5.3 Optical Tweezers

Optical tweezers form another important tool for single-molecule studies of proteins. Basically, optical tweezers utilize the force of laser radiation pressure to trap small particles. Following this, other associated techniques are used to manipulate the trapped particles and study the variation of forces (Novotny et al. 1997; Curtis et al. 2002). Initially, this methodology was used for manipulating and studying micron-sized dielectric particles. Nevertheless, now optical tweezers have been implemented in studying biological systems to a good extent. Mainly, the ability to manipulate single molecules at nanometre scale and to measure forces on these molecules has made this technique an important biophysical tool. In general, the radiation pressure is defined as a force per unit area on a molecule due to change in the momentum of light. All light is composed of photons having a consistent momentum (p). And if there is a change in momentum of light due to refraction with a molecule, the difference between the momentum flux entering and leaving the object can be used to define the total force. Thus, force due to refraction of light can be used to trap and manipulate molecules and helps in making quantitative measurements (Smith et al. 2003). Conformational changes in molecules can generate change in these measurements and thus can provide opportunity to demonstrating their conformational space under particular physicochemical condition. In other words, these measurements can be used to study structural transitions in molecules, define presence or absence of intermediates or even predict the presence of an alternate pathway. Moreover, the wealth of kinetic and thermodynamic information can be obtained that include rates of transitions, free-energy calculations, identification and calculation of energy barriers and even the full energy landscape showing the conformational dynamics of a protein (Wang et al. 1998).

Recently, many workers have used optical tweezers to obtain insights into the conformational paradigm of intrinsically disordered proteins (IDPs). The technique has helped in probing the low-energy fluctuations and marginally stable structures of IDPs. A very important example is the study of α-synuclein, an IDP that is associated with the pathophysiology of Parkinson's disease (Neupane et al. 2014). Several transient metastable structures of this protein have been identified and studied using optical tweezers. This has helped in showing quasi-equilibrium fluctuations in α-synuclein structure that are short-lived and very unstable that may be attributed to the previously predicted molten globule-like conformation. Besides, several groups have also studied transient states of proteins in their heterogeneous mixtures such as during misfolding and aggregation (Stangner et al. 2013; Heidarsson et al. 2014;

Mashaghi et al. 2014). Recent work on dimeric prion protein has shown the presence of multiple intermediates with comparatively slower diffusion rates that indicates a rugged or rough energy landscape of misfolding (Yu et al. 2015). Similarly, other studies involving optical tweezers have shown the presence of misfolded intermediates in case of slow refolding luciferase enzyme and have helped in correlating calcium concentration and misfolding of neuronal calcium sensor-1 (NCS-1) (Heidarsson et al. 2014). Overall, optical tweezers form an interesting tool to study the dynamics and function of proteins under both physiological and non-physiological conditions.

2.6 Proteases in Biotechnology

Proteases form one of the largest family of enzymes that constitute about 2% of the human genome and account for more than 60% of the total enzyme market. In general, proteases are enzymes that perform proteolysis in a specific or non-specific manner. These proteins mainly hydrolyse peptide bonds and are commonly found associated with protein catabolic pathways. These catabolic pathways in turn form a part of variety of important physiological processes such as cell division, apoptosis, signal transduction, food digestion, blood clotting, etc. Besides, several disease aetiologies are found associated with variety of proteases playing a part in growth and progression of pathogens including the replication of viruses (Neurath and Walsh 1976). Thus, these proteins harbour a diversity of structural and functional features that help them carry out critical purposes inside and outside the cells. Despite several evolutionarily linked diversifications, all proteases essentially function to cleave the amide bond by a nucleophilic attack on the carbonyl group that results in hydrolysis. These proteases naturally exist as inactive precursors called zymogens. These zymogens get converted to their active forms by particular environmental triggers. These triggers could be a variety of factors such as pH, salt, ions, small molecules, peptides or specific ligands that induce specific conformational changes in the protein (Anwar and Saleemuddin 1998; Gupta et al. 2002).

Owing to the huge demand of industrial proteases, sources other than plants and animals have garnered much interest in the scientific and industrial community. This increased interest has led to the discovery of numerous microbial proteases as an alternative source. Basically, the broad biochemical diversity of microbial enzymes as well as their high yield has led to their high acceptability and success at the commercial scale. Additionally, protein engineering approaches to modify the cleavage specificities, solubility and stability of proteases have further broadened their commercial landscape (Eijsink et al. 2004; Bansal et al. 2012). These engineered next-generation proteases are tailored to increase substrate selectivity as well as lend multiple substrate hydrolyzing capabilities. Besides other industrial applications, proteases are also advantageous therapeutic agents that have been employed in the treatment of several diseases such as cancer, AIDS, cardiovascular disease and other infections (Rao et al. 1998; Chanalia et al. 2011).

2.6.1 Classification of Proteases

Proteases are broadly classified based on their site of action and by the nature of active site residues. The Enzyme Commission (EC) classifies proteases into group 3 (hydrolases) and subgroup 4 (peptide bond hydrolases). On the basis of site of action, proteases are further categorized as endopeptidases and exopeptidases. The exopeptidases specifically attack on the peptide bonds proximal to the start or end of protein substrate, i.e. either at the amino- or carboxy-terminal. Thus, based on their N- or C-terminus specificity, these enzymes are further divided into aminopeptidases or carboxypeptidases. Besides, the endopeptidases cleave peptide bonds inside the polypeptide chain mainly away from the C- or N-termini.

Further, these endopeptidases are classified into six major groups based on the nature of active site residues involved in their catalytic mechanism (Garcia-Carreon and Del Toro 1997). These are called aspartate, glutamate, cysteine, serine, threonine and metalloproteases. The serine, cysteine and threonine proteases involve nucleophilic attack on the substrate and produce covalently linked acyl-enzyme intermediates. On the other hand, aspartate, glutamate and other metalloproteases activate water molecules that act as nucleophile for breaking the peptide bond. Additionally, based on pH optima, proteases are also categorized as acidic, neutral or alkaline proteases.

2.6.2 Protease Sources

Proteases are ubiquitously found in all forms of life and play a variety of physiological functions. Owing to their large presence in the living milieu, initially, these proteases were commercially obtained from several plant and animal sources directly. However, due to low agricultural facilities specifically for industrial production of proteases, plant sources are limited. Similarly, animal origin proteases such as trypsin, pepsin and rennins are extracted directly in bulk quantities from animal sources (Gupta et al. 2002). Nevertheless, their production also depends on livestock availability and ethical and operational limitations. Thus, low production from plant and animal sources and increasing industrial demands led to identification of other alternative sources of proteases (Table 2.3).

Thus, owing to their huge biochemical diversity, microbial sources are slowly becoming commercial generators of various proteases. With the advent of new protein engineering approaches, it has become easier to optimize and regulate microbial growth and hence enzyme production. Besides the ease of bulk production, the other major advantage of using microbial enzymes is their similar characteristics compared to original animal and plant enzymes.

2.6.3 Engineered Proteases

The proteases have been commercially utilized in several applications and are often required to work under extreme and non-physiological conditions such as high temperature, pH, presence of chaotropes, chelating agents and detergents (Gupta et al.

Table 2.3 Few important commercial proteases and their source of production

Product	Company	Protease type	Source
Protex 30L	Genencor/DuPont	Serine protease	*Bacillus subtilis*
Neutrase 0.8L	Novozymes	Protease	*B. amyloliquefaciens*
Protex 6L	Genencor/DuPont	Serine protease	*B. amyloliquefaciens*
Ronozyme ProAct	DSM/Novozymes	Alkaline serine protease	*B. licheniformis*
Esperase 8.0L	Novozymes	Protease	*Bacillus* spp.
FNA	Genencor/DuPont	Serine protease	*B. amyloliquefaciens*
Versazyme	Novus	Keratinase	*B. licheniformis*
Arazyme-One-Q	Insect Biotech Co.	Metalloprotease	*Serratia proteamaculans* HY-3
Alcalase 2.4	Novozymes	Subtilisin	*Bacillus* spp.
Savinase	Novozymes	Subtilisin	*Bacillus* spp.
Porperase 1600L	Genencor/DuPont	Serine protease	*B. alcalophilus*
Allzyme FD	Alltech	Serine protease	*Aspergillus niger*

2002). This leads to lowered efficacies and degradation of proteases in industrial setups. Thus, proteases with improved stability and functions have been produced using modern-day genetic engineering, protein engineering and directed evolution approaches. Beside this, the proteases have also been used as therapeutics as they are capable of cleaving various physiological substrates. However, very few proteases have been used as therapeutics because of narrow substrate specificity, stability and activity. The development of high-throughput, bacterial and yeast-based methods has facilitated the redesigning of proteases to the modified one with novel specificities, lesser toxicities and higher stability (Guerrero et al. 2016).

In leather and detergent industries, proteases with high thermostability are required for optimal activity and yield. In general, high temperatures affect the native fold of proteins leading to aggregation and loss of activity. Protease engineering approaches have been successful in enhancing thermostability of proteases to a good extent (Mitrea et al. 2010; Zhao et al. 2016). These approaches mainly rely on availability of thermostable homologue of particular enzyme/protease that is used to identify crucial residues and interactions. These critical residues are then engineered into the mesophilic protein scaffolds to mimic their thermostable counterparts. Another widely used approach is the introduction of non-natural disulphides in the protein structures to confer more stability. The addition of extra-stabilizing disulphides helps in lowering the entropy of the protein and thus keeping the native and active scaffold of the protein intact. An important example is the production of engineered thermostable subtilisin E that incorporates an unnatural disulphide based on structural comparisons with a thermophilic subtilisin-type protease, aqualysin I (Bryan 2000).

Similarly, metal binding sites in many enzymatic proteins have been found to stabilize and help increase the activity. Introduction of calcium-binding pockets by substituting Gly131 and Pro172 by Asp residues in subtilisin increased electrostatic interactions and stabilized the protein (Pantoliano et al. 1988). On the contrary, industries such as detergent and laundries require proteins to function under highly chelating conditions. Thus, proteins with high stability aided with better catalytic function in absence of metal ligands have been produced by a combination of

protein engineering and directed evolution. A subtilisin variant Sbt88 with 1000 times greater stability was characterized based on deletion of calcium-binding loop followed by improvement of activity by directed evolution (Strausberg et al. 1995). Also there are several industrial applications that require proteases to function in organic solvents. Several methodologies that mainly include error-prone PCR and directed evolution have been utilized to screen proteases with improved activity in organic solvents.

Besides other industrial applications, proteases have also been extensively used in wide range of therapeutics and have been engineered to suit particular application. Engineered factor VIIa (FVIIa) with improved pro-coagulant and anti-fibrinolytic activity by residue mutations and increasing its enzymatic efficacy by increased localization affinity are such few examples (Persson et al. 2001). Another important aspect of engineering therapeutic proteases has been to reduce their non-specificity and thus the side effects. On these lines, bifunctional thrombin represents an excellent example where the protein not just acts as an activator of multiple pro-coagulant substrates but also acts as an anticoagulant agent (Hanson et al. 1993). Further, thrombin has also been engineered to produce a more efficient protein having higher anticoagulant function. An engineered double mutant thrombin (W215A/E217A) with compromised pro-coagulant activity and very slow (>2000-fold) fibrinogen cleaving activity has been developed that specifically activates protein C in the presence of thrombomodulin (Gandhi et al. 2009). Another important factor associated with therapeutic proteases or proteins in general is their half-lives under physiological milieu. Several factors such as protease digestion and endogenous inhibition have been attributed for their short half-lives in organismic environment. Among several examples, designing inhibitors of tissue plasminogen activator (t-PA) has been shown to prevent systemic activation of plasminogen and helped in promoting thrombolytic property of t-PA (Madison et al. 1989). Furthermore, reducing immunogenicity of therapeutic proteases has also been a subject of interest to increase their overall clinical efficacy. Thus, novel protein engineering approaches along with modern-day computational biochemistry and synthetic biology have helped in channelizing the development of better and high-quality proteases with improved commercial and industrial applications (Table 2.3).

2.6.4 Production Process

The animal- and plant-based enzymes are directly produced from the animal tissues and extracts. Initially, the tissue and organs are processed in an enzyme-extracting solution to obtain the crude proteases. Later, several steps of extraction and concentration are involved to get proteases of typical potency (Jaswal et al. 2008). Nowadays, most major proteases and enzymes are obtained at commercial scales using microbial and fungal strains. These protease-producing strains are screened using plate assay method, and their production capability is estimated quantitatively. At the next stage, media design and optimization are done to help in inducing the production of particular enzyme/protease. The optimization process involves

find application in other basic science protocols such as general protein digestion, removal of affinity tags, tissue extraction and separation and several cell culture methodologies. The rise in engineered proteases in recent times has brought down the cost of production in major industries such as pharmaceuticals, bakery, detergent, leather and food. Protein engineering helps in incorporating desirable characteristics such as heat stability, pH range, substrate specificity, increased activity and storage time. Proteases from various sources such as bacterial, fungal and viral have been engineered to produce suitable application-based enzymes.

2.7 Chaperones in Biotechnology

Molecular chaperones have been an area of growing interest in the field of medical, pharmaceutical and industrial biotechnology (Henderson et al. 1996; Arakawa et al. 2006). Chaperone function has been found critical in both normal and stress conditions in all living cells. The term chaperone dates back to 1978 where it was used to describe the function of a nuclear protein called nucleoplasmin that helps in chromatin assembly by preventing nonspecific interaction of histone proteins with DNA (Schlieker et al. 2002). Later, these were categorized as a class of proteins that assist in post-translational modification of other proteins. The current understanding of chaperones describes them as a highly conserved protein family that primarily functions to assist folding of nascent polypeptides and prevent their misfolding and aggregation (Hartl and Hayer-Hartl 2002). This stabilizing function of chaperones has been responsible for protecting proteins under variety of stress conditions including heat, chemicals, toxins, free radicals, radiations, etc. Chaperones binding to exposed hydrophobic sites in polypeptides help in preserving their correct scaffold thereby preventing their aggregation. Their role in aggregation prevention has been probed and studied in several neurodegenerative and systemic disorders such as Alzheimer's, Parkinson's, diabetes, etc. (Henderson et al. 1996; Barral et al. 2004). Besides, the role and application of chaperones in folding proteins of therapeutic and industrial importance have been areas of intense research.

The chaperoning function is not just associated with protein-based chaperones but has also been successfully achieved using certain chemical scaffolds and co-solvents too. These co-solvents include several osmolytes and small molecules such as arginine that help in raising the osmotic pressure against environmental water stresses and result in protein stabilization (Arya et al. 2014; Dandage et al. 2015; Srivastava et al. 2015). The well-characterized genetic architecture as well as the ease of fermentation has made *E. coli* as an important industrial organism for the production of proteins of various origins. It has been shown that chaperones can help in improving the refolding process obtained from inclusion bodies and may also be expressed in *E. coli* to reduce the formation of inclusion bodies (Thomas and Baneyx 1997). Thus, in vivo overexpression of chaperones along with protein of interest has been utilized as a successful strategy to obtain proteins of industrial and commercial interest with high yields (Villaverde and Carrio 2003; Ventura and Villaverde 2006). Besides, there had been several other strategies where the yields

of industrial proteins have been raised by addition of chemical chaperones at different steps of purification process (Mannen et al. 2001; Rajan et al. 2011).

Moreover, several studies have also demonstrated that chaperones coexist and cooperate with each other during the refolding process in vivo (Hartl and Hayer-Hartl 2002; Genevaux et al. 2007). Hence, there has been a rise in newer strategies that involve expression of a combination of chaperones rather than a single chaperone. However, there are several limitations associated with chaperone overexpression, and hence including more steps in the in vitro refolding and purification process has been favoured in the recent years (Hoffmann and Rinas 2004; de Marco et al. 2007). This has led to utilization of protein chaperones as well as chemical and pharmacological chaperones in the industrial setups. Besides, it has been also found that immobilization of chaperones has been also identified as an important tool for retaining and reutilization of chaperones for longer usage. As a whole, chaperones have been identified as crucial tools for improving recombinant production of therapeutic and industrial important proteins that are aggregation prone.

2.7.1 Chaperone-Assisted Protein Folding

Molecular chaperones bind selectively to nascent polypeptides and partially folded protein intermediates and help them fold into their native and active structure. Thus, chaperones effectively demarcate folded and unfolded polypeptides based on their hydrophobic features (McHaourab et al. 2009). Although the primary sequence contains inherent information on protein folding, the crowded cellular milieu makes a nascent polypeptide very prone to misfolding and aggregation. At this stage, chaperones play a crucial role in stabilizing and assisting the initial protein scaffold to correctly fold into its active form. Besides, molecular chaperones are often involved in compartmental translocations, proteolytic degradation and removal of misfolded and aggregation-prone proteins (Young et al. 2003; Young et al. 2004).

In general, chaperones are found in both prokaryotic and eukaryotic systems and are similar in their basic structural features. The major differences occur in quaternary organization involving the number of subunit, monomer associations and their active oligomeric states. Nevertheless, all chaperones function in a similar fashion and bind to aggregation prone and partially folded protein intermediates to prevent their aggregation. Out of several identified systems, the best characterized chaperone system considers a chamber-based folding of a nascent polypeptide. The GroEL-GroES chaperonin system represents the best example of this chamber-based protein folding (Mayhew et al. 1996). Here, the GroES molecule acts as the lid of the folding chamber in the form of GroEL and helps in creating an isolated environment for a nascent polypeptide. The folding is assisted in the presence of ATP, and the phenomenon is explained via *cis*- or *trans*-mechanism (Hayer-Hartl et al. 2015). The folding mechanism in most chaperone systems is ATP dependent and is considered highly efficient and error-free. The protein folding funnel describes large number of protein conformations with very high entropy at the top. During the folding process, the entropy slowly decreases owing to the number of unstable states receding into native or near-native conformations (Baldwin and Rose 2013).

Here, it is worth mentioning that smaller proteins generally undergo spontaneous folding and large proteins mostly require chaperone assistance for correct folding.

Chaperone assistance actually helps in reducing the population of unfolded conformers that represent different energy levels. Thus, a chaperone-assisted channelization of the energy funnel helps the protein population to attain native structure. In most cases, along with the main chaperone, several co-chaperones smaller in sizes are also present. A major example is the DnaK chaperone system that involves formation of ATP-bound DnaJ complex with DnaJ co-chaperonin (Rosenzweig et al. 2013). The process follows ATP hydrolysis by water and addition of another subunit GprE that releases both ADP and DnaJ. This propagates the release of folded protein from the chaperone complex. Besides these in vivo chaperone systems that have been utilized for recombinant protein production, several chemicals and solvents have been also used to refold proteins in vitro (Bruzdziak et al. 2013). This group of chemical chaperone consists mainly of osmolytes and non-osmolytic small molecules. Chemical chaperones also function by sticking and stabilizing the exposed hydrophobic patches of partially folded proteins and hence facilitate their folding (Rajan et al. 2011). Thus, they form a key step in obtaining correctly refolded proteins in industrial setups. Osmolytes augment the thermodynamic stability and thus aid in protein folding. Besides, solvent additives like arginine also help in recovering active protein during large-scale production of bacterially expressed proteins.

2.7.2 Classes of Chaperones

Molecular chaperones are categorized into two major classes, viz. group I type chaperones found in bacteria, mitochondria and plastids and the group II type chaperones found only in archaeal and eukaryotic cytoplasm (Johnson and Craig 1997). Besides, there is another group of co-chaperones that assist chaperones in selecting protein targets and regulate their association and dissociation. Based on function and mechanism of action, molecular chaperones have been further classified into three subclasses. The 'folding' chaperones (e.g. DnaK and GroEL) depend on ATP-driven conformation change for polypeptide substrate refolding. The 'holding' chaperones (e.g. Hsp 31 and IbpA) bind to the partially unfolded protein intermediates until the 'folding' chaperones are available during stress conditions. Besides, the third class of chaperones (e.g. ClpB) acts on already aggregated forms of proteins and helps their resolubilization. On the other hand, there are two separate classes of nonbiological chaperones that include 'chemical chaperones' consisting of mainly osmolytes (e.g. trehalose) and 'pharmacological chaperones' that include mainly protein-stabilizing ligands and small molecules (e.g. arginine).

2.7.3 Aspects in Industrial Protein Production

The well-characterized genetics as well as fermentation technology of *E. coli* has made it the most commonly used organism for the heterologous production of proteins from various sources. However, the major bottleneck with *E. coli*-based

Table 2.4 Some important chaperones expressed in specific host cells for assisting protein folding

Chaperone	Host
GroEL/ES	*E. coli*
BiP	*Insect cells*
DegP	*E. coli*
Peptidylprolyl *cis-trans* isomerise	*Insect cells*
DnaK	*E. coli*
ClpB	*E. coli*
Hsp 70	*Insect cells*
HtbG	*E. coli*
Calnexin	*Insect cells*
Human protein disulphide-isomerase (PDI)	*CHO cells*
Polyubiquitin	*Kluyveromyces lactis*
DegP	*E. coli*

protein production is the formation of insoluble inclusion bodies (IB). Although IBs help in greatly increasing the overall yield of the recombinant protein, the process of obtaining the protein in its function form is often problematic (Ventura and Villaverde 2006). The exact mechanism behind the formation of these depositions is not known, but it is assumed that eukaryotic nascent polypeptides generally have higher exposure of hydrophobic sites and thus have higher affinity for aggregation (Fahnert et al. 2004). The other important factor is the inefficient protein chaperoning and stabilization system in the bacteria. This led to large number of reports supporting the usage of chaperone overexpression as an immediate solution to this problem (Georgiou and Valax 1996; Thomas and Baneyx 1996; Carrio and Villaverde 2002) (Table 2.4).

Molecular chaperones like GroEL/ES or DnaK have been co-expressed with desired proteins to obtain considerable yields (Chen et al. 2003; Gupta et al. 2006). Nevertheless, it was found that expression of single chaperones does not have significant increase in the amount of soluble protein. Further, it was found that large chaperones actually require small co-chaperones that function in consort to 'hold' and 'refold' the protein. It was shown that application of bi-chaperone systems such as ClpB and DnaK could be a better resort in obtaining soluble protein in significantly higher yield as compared to their individual co-expression (Mogk et al. 1999; Mogk et al. 2003). The eukaryotic counterparts of both these proteins (Hsp70 and Hsp40) have been also shown to improve refolding of recombinant proteins as compared to co-expression of large chaperones such as Hsp90 alone (Glover and Lindquist 1998; Genest et al. 2013). Molecular chaperones are currently utilized in biotechnology industry for their ability to prevent as well as reverse the protein aggregation. Several complex systems such as the *E. coli* DnaK-DnaJ-GrpE system have been utilized to suppress protein aggregation and effectively perform disaggregation (Zolkiewski 1999).

On the other hand, owing to their complex mechanism of action, the precise calculation of aggregation kinetics as well as affinities of chaperone systems (bi-, tri- or more) to different aggregate types are not yet characterized. This fall back has led to several model propositions for defining the actual phenomenon. Two major

models are widely accepted that explain the mechanism of action of multiple chaperone systems. The DnaK/ClpB system (chaperone/co-chaperone system) has been studied in detail and has been utilized to develop these models (Mogk et al. 1999). According to the first model, ClpB makes initial contacts with the protein aggregates and later gets attached to DnaK. This model is based on the finding that DnaK has relatively lower binding affinity of binding to smaller protein aggregates. On the other hand, the second model suggests that the ClpB/DnaK complex directly attaches with the protein which is solely determined by specificity of individual chaperone components. Nevertheless, at present, there is still a dearth of significantly active chaperone systems that could be applied in the industrial setups effectively.

2.8 Therapeutic Proteins

There has been a surge in protein-based therapeutics during the past few decades. The major reason behind this surge is the critical role of proteins in cellular machinery and elucidation of mechanisms of their function in great detail. This led to the development of an entirely new class of biotherapeutics that is produced from living systems or their products (Carter 2011). Protein-based biotherapeutics extensively accounts for monoclonal antibodies, recombinant proteins, fusion proteins and peptides. Owing to their broad-spectrum specifications, therapeutic proteins have been successfully utilized in treating cancers, cardiovascular diseases, diabetes, anaemia, haemophilia, etc. (Leader et al. 2008). Nevertheless, these proteins have their own demerits and limitations such as low oral stability due to enzymatic degradation, low physiological half-life and poor shelf life. Thus, the area of protein therapeutics is continuously being researched upon to make them better and more efficient. Next-generation protein therapeutics relies heavily upon protein engineering that helps in modifying these proteins of different origins to suit human immunological response. Currently, these emerging new pharmaceutical products account for more than 200 commercial products that mainly include therapeutics, diagnostics and vaccines (Pavlou and Reichert 2004) (Table 2.5).

2.8.1 Classes and Source

Protein therapeutics have been broadly classified based on their pharmacological activity. They can be used as replacement to abnormal or deficient protein and as delivery systems for certain drugs, effectors or toxins, for enhancing a particular metabolic pathway, adding a new function or activity, and interfere or inhibit certain pathological molecule or an organism (Caravella and Lugovskoy 2010). Besides this function-based categorization, they have been also grouped based on molecular type such as blood factor, antibody, anticoagulant, enzyme, scaffold, fusion protein, growth factors, etc. Overall, protein-based therapeutics is an ever-growing field that relies on producing newer, engineered, efficient and immunologically less-reactive protein-based pharmaceuticals.

protein-based therapeutics and has greatly revolutionized their clinical usage. Engineered natural proteins, fusion proteins and PEG conjugated and antibody-based therapeutics are only few examples that have been widely utilized for producing effective therapeutics (Carter 2011). Insulin is the first therapeutic protein produced through recombinant DNA technology that was commercialized. This synthetic counterpart (Humulin, Eli Lilly Inc.) was approved by the US Food and Drug Administration (FDA) in 1982 for the treatment of diabetes mellitus (Johnson 1983). The two major advantages were its rapid action after meals and also maintenance of low levels of insulin between meals. The engineered insulin incorporating stabilizing amino acid mutations was similar in action to the endogenous insulin in the body. Later, several companies modified and produced more efficient insulin analogues that had better physiological half-lives, reduced risk of hyperglycaemia and ease of administration. These rationalizations were employed in commercialization of several other natural proteins such as erythropoietin and tissue plasminogen activator (tPA) and different cytokines such as interferon-α (Andersen and Krummen 2002).

Several other protein-based therapeutics have been produced in mammalian cells with additional glycosylation sites that help in augmenting the in vivo half-life. For example, engineered darbepoetin (Aranesp) is administered to anaemic patients for enhancing the production of RBCs (Egrie and Browne 2002). Similarly, epoetin-α has been also produced to contain extra N-linked glycosylation sites for better physiological half-life. Besides this, the advent of fusion proteins has opened a new era of protein-based therapeutics. Fusion proteins are created to combine properties of two or more different proteins and have been commercially successful in countering many diseases such as cancer and multiple sclerosis (Dingermann 2008). An excellent example is the fusion of Fc region of immunoglobulins to target proteins that help in increasing their serum retention and half-life. Apart from increasing half-lives, conjugation with Fc region enables high-throughput purification and increased solubility and stability of therapeutic proteins. Important examples of Fc-fusion proteins that have been approved after clinical trials include etanercept (Enbrel) incorporating a TNF receptor 2 and romiplostim, a thrombopoietin (TPO) receptor agonist that is composed of peptide-Fc fusion (Bibila and Robinson 1995; Chapman et al. 1999).

Other than this, polyethylene glycol (PEG) conjugation has also been shown to improve clinical efficacy of therapeutic proteins. This has given rise to an entirely new engineering technology called 'PEGylation' that has been employed in improving efficacy of several proteins (Ginn et al. 2014). PEGylated proteins have higher molecular volume that prevents their early renal clearance leading to an increased half-life. Several PEGylated proteins have been commercialized successfully as efficient therapeutic proteins that include analogues of interferon-α, interferon-β1a, erythropoietin, human GCSF (granulocyte colony-stimulating factor) and L-asparaginase, etc.

Moreover, there has been a rise in production of 'engineered protein scaffolds' that act as alternative to antibody-based therapeutics since they recognize specific protein targets (Gebauer and Skerra 2009). These scaffolds are produced as a library

of small and highly soluble moieties and later screened using phage or yeast systems against specific protein targets. A number of these scaffolds are now increasingly adding up into the next-generation therapeutics (Skerra 2000; Lofblom et al. 2011). Several examples include avimer (based on LDLR-A modules), Adnectins (based on type III fibronectin domains), affibodies (based on synthetic Z domain of staphylococcal protein) and DARPins (based on ankyrin repeat proteins). These scaffolds owing to their small sizes are now utilized for creating multi-targeting and multivalent next-generation protein-based therapeutics. Thus, the clinical success of protein-based therapeutics has revolutionized the clinical development of several new, engineered and efficient protein drugs against several pernicious diseases.

Conclusion

The postgenomic era has paved new pathways into the discovery and development of novel proteins and protein/enzyme-based products. This huge information base has aided protein engineering and later bioprocess to emerge as more efficient and product oriented. The molecular details of protein functions have been elucidated using modern-day single-molecule studies. Utilizing this functional information along with structural biology, several new modifications in industrial protein production have been implemented. In this context, traditional methodologies including directed evolution approaches are now aided with highly curated sequence information as well as more precise molecular biology tools. Besides, the industrial bioprocess setup for commercial proteins/enzymes now includes chaperone-based methodologies. Both at the in vitro as well as in vivo levels, protein recovery is greatly enhanced using biological and chemical chaperones. Overall, the present scenario in protein and enzyme production at industrial scale looks highly competitive and exacting due to increasingly high stimulus of information on sequence, structural and functional aspects of proteins.

References

Acevedo-Rocha CG, Reetz MT (2014) Assembly of designed oligonucleotides: a useful tool in synthetic biology for creating high-quality combinatorial DNA libraries. Methods Mol Biol 1179:189–206. doi:10.1007/978-1-4939-1053-3_13

Ahmad P, Ashraf M, Younis M, Hu X, Kumar A, Akram NA, Al-Qurainy F (2012) Role of transgenic plants in agriculture and biopharming. Biotechnol Adv 30:524–540. doi:10.1016/j.biotechadv.2011.09.006

Anczykowski B, Gotsmann B, Fuchs H, Cleveland JP, Elings VP (1999) How to measure energy dissipation in dynamic mode atomic force microscopy. Appl Surf Sci 140:376–382. doi:10.1016/S0169-4332(98)00558-3

Andersen DC, Krummen L (2002) Recombinant protein expression for therapeutic applications. Curr Opin Biotechnol 13:117–123. doi:10.1016/S0958-1669(02)00300-2

Anwar A, Saleemuddin M (1998) Alkaline proteases: a review. Bioresour Technol 64:175–183. doi:10.1016/S0960-8524(97)00182-X

Arakawa T, Ejima D, Kita Y, Tsumoto K (2006) Small molecule pharmacological chaperones: from thermodynamic stabilization to pharmaceutical drugs. Biochim Biophys Acta 1764:1677–1687. doi:10.1016/j.bbapap.2006.08.012

Arya P, Srivastava A, Vasaikar SV, Mukherjee G, Mishra P, Kundu B (2014) Selective interception of gelsolin amyloidogenic stretch results in conformationally distinct aggregates with reduced toxicity. ACS Chem Neurosci 5:982–992. doi:10.1021/cn500002v

Baldwin RL, Rose GD (2013) Molten globules, entropy-driven conformational change and protein folding. Curr Opin Struct Biol 23:4–10. doi:10.1016/j.sbi.2012.11.004

Bansal S, Srivastava A, Mukherjee G, Pandey R, Verma AK, Mishra P, Kundu B (2012) Hyperthermophilic asparaginase mutants with enhanced substrate affinity and antineoplastic activity: structural insights on their mechanism of action. FASEB J 26:1161–1171. doi:10.1096/fj.11-191254

Barral JM, Broadley SA, Schaffar G, Hartl FU (2004) Roles of molecular chaperones in protein misfolding diseases. Semin Cell Dev Biol 15:17–29. doi:10.1016/j.semcdb.2003.12.010

Best RB, Hofmann H, Nettels D, Schuler B (2015) Quantitative interpretation of FRET experiments via molecular simulation: force field and validation. Biophys J 108:2721–2731. doi:10.1016/j.bpj.2015.04.038

Besta RB, Brockwellb DJ, Toca-Herreraa JL, Blakeb AW, Smithc AD, Radford SE, Clarke J (2003) Force mode atomic force microscopy as a tool for protein folding studies. Anal Chim Acta 479:87–105. doi:10.1016/S0003-2670(02)01572-6

Bibila TA, Robinson DK (1995) In pursuit of the optimal fed-batch process for monoclonal antibody production. Biotechnol Prog 11:1–13. doi:10.1021/bp00031a001

Bruzdziak P, Panuszko A, Stangret J (2013) Influence of osmolytes on protein and water structure: a step to understanding the mechanism of protein stabilization. J Phys Chem B 117:11502–11508. doi:10.1021/jp404780c

Bryan PN (2000) Protein engineering of subtilisin. Biochim Biophys Acta 1543:203–222. doi:10.1016/S0167-4838(00)00235-1

Buiatti M, Christou P, Pastore G (2013) The application of GMOs in agriculture and in food production for a better nutrition: two different scientific points of view. Genes Nutr 8:255–270. doi:10.1007/s12263-012-0316-4

Butt HJ, Cappella B, Kappl M (2005) Force measurements with the atomic force microscope: technique, interpretation and applications. Surf Sci Rep 59:1–152. doi:10.1016/j.surfrep.2005.08.003

Bylund F, Collet E, Enfors SO, Larsson G (1998) Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. Bioprocess Eng 18:171–180. doi:10.1007/s004490050427

Camps M, Naukkarinen J, Johnson BP, Loeb LA (2003) Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. Proc Natl Acad Sci U S A 100:9727–9732. doi:10.1073/pnas.1333928100

Caravella J, Lugovskoy A (2010) Design of next-generation protein therapeutics. Curr Opin Chem Biol 14:520–528. doi:10.1016/j.cbpa.2010.06.175

Carrio MM, Villaverde A (2002) Construction and deconstruction of bacterial inclusion bodies. J Biotechnol 96:3–12. doi:10.1016/S0168-1656(02)00032-9

Carter PJ (2011) Introduction to current and future protein therapeutics: a protein engineering perspective. Exp Cell Res 317:1261–1269. doi:10.1016/j.yexcr.2011.02.013

Carvalho FA, Santos NC (2012) Atomic force microscopy-based force spectroscopy—biological and biomedical applications. IUBMB Life 64:465–472. doi:10.1002/iub.1037

Celik E, Calık P (2012) Production of recombinant proteins by yeast cells. Biotechnol Adv 30:1108–1118. doi:10.1016/j.biotechadv.2011.09.011

Chanalia P, Gandhi D, Singh JD (2011) Applications of microbial proteases in pharmaceutical industry: an overview. Rev Med Microbiol 22:96–101. doi:10.1097/MRM.0b013e3283494749

Chapman AP, Antoniw P, Spitali M, West S, Stephens S, King DJ (1999) Therapeutic antibody fragments with prolonged in vivo half-lives. Nat Biotechnol 17:780–783. doi:10.1038/11717

Chen R (2012) Bacterial expression systems for recombinant protein production: *E. coli* and beyond. Biotechnol Adv 30:1102–1107. doi:10.1016/j.biotechadv.2011.09.013

Chen I, Ting AY (2005) Site-specific labelling of proteins with small molecules in live cells. Curr Opin Biotechnol 16(1):35–40

Chen Y, Song J, Sui SF, Wang DN (2003) DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA over expressed in *Escherichia coli*. Protein Expr Purif 32:221–231. doi:10.1016/S1046-5928(03)00233-X

Cheong C, Chun SS, Kim YH (1993) Production and properties of an alkaline protease from *Pseudomonas sp.* SJ-320. Korean Biochem J 26:479–484

Churnside AB, Sullan RM, Nguyen DM, Case SO, Bull MS, King GM, Perkins TT (2012) Routine and timely sub-picoNewton force stability and precision for biological applications of atomic force microscopy. Nano Lett 12:3557–3561. doi:10.1021/nl301166w

Clark ED (2001) Protein refolding for industrial processes. Curr Opin Biotechnol 12:202–207. doi:10.1016/S0958-1669(00)00200-7

Cobb RE, Si T, Zhao H (2012) Directed evolution: an evolving and enabling synthetic biology tool. Curr Opin Chem Biol 16:285–291. doi:10.1016/j.cbpa.2012.05.186

Coco WM, Levinson WE, Crist MJ, Hektor HJ, Darzins A, Pienkos PT, Squires CH, Monticello DJ (2001) DNA shuffling method for generating highly recombined genes and evolved enzymes. Nat Biotechnol 19:354–359. doi:10.1038/86744

Craik CS, Page MJ, Madison EL (2011) Proteases as therapeutics. Biochem J 435:1–16. doi:10.1042/BJ20100965

Cramer CL, Boothe JG, Oishi KK (1999) Transgenic plants for therapeutic proteins: linking upstream and downstream strategies. Curr Top Microbiol Immunol 240:95–118. doi:10.1007/978-3-642-60234-4_5

Crameri A, Raillard SA, Bermudez E, Stemmer WP (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288–291. doi:10.1038/34663

Curtis JE, Koss BA, Grier DG (2002) Dynamic holographic optical tweezers. Opt Commun 207:169–175. doi:10.1016/S0030-4018(02)01524-9

Dandage R, Bandyopadhyay A, Jayaraj GG, Saxena K, Dalal V, Das A, Chakraborty K (2015) Classification of chemical chaperones based on their effect on protein folding landscapes. ACS Chem Biol 10:813–820. doi:10.1021/cb500798y

de Marco A, Deuerling E, Mogk A, Tomoyasu T, Bukau B (2007) Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. BMC Biotechnol 12:7–32. doi:10.1186/1472-6750-7-32

Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv 27:297–306. doi:10.1016/j.biotechadv.2009.01.008

den Engelsman J, Garidel P, Smulders R, Koll H, Smith B, Bassarab S, Seidl A, Hainzl O, Jiskoot W (2011) Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. Pharm Res 28:920–933. doi:10.1007/s11095-010-0297-1

Deniz AA, Laurence TA, Beligere GS, Dahan M, Martin AB, Chemla DS, Dawson PE, Schultz PG, Weiss S (2000) Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. Proc Natl Acad Sci U S A 97:5179–5184. doi:10.1073/pnas.090104997

Deniz AA, Mukhopadhyay S, Lemke EA (2008) Single-molecule biophysics: at the interface of biology, physics and chemistry. J R Soc Interface 5:15–45. doi:10.1098/rsif.2007.1021

Desai KM, Survase SA, Saudagar PS, Lele SS, Singhal RS (2008) Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. Biochem Eng J 41:266–273. doi:10.1016/j.bej.2008.05.009

Dingermann T (2008) Recombinant therapeutic proteins: production platforms and challenges. Biotechnol J 3:90–97. doi:10.1002/biot.200700214

Doran PM (2000) Foreign protein production in plant tissue cultures. Curr Opin Biotechnol 11:199–204. doi:10.1016/S0958-1669(00)00086-0

Dufrene YF (2002) Atomic force microscopy, a powerful tool in microbiology. J Bacteriol 184:5205–5213. doi:10.1128/JB.184.19.5205-5213.2002

Egrie JC, Browne JK (2002) Development and characterization of darbepoetin alfa. Oncology (Williston Park) 16:13–22

Eijsink VGH, Bjorkb A, Gaseidnesa S, Sirevagb R, Synstada B, van den Burgc B, Vriendd G (2004) Rational engineering of enzyme stability. J Biotechnol 113:105–120. doi:10.1016/j.jbiotec.2004.03.026

Elleke FB, van der Oost J, Vos MW, van Kranenburg WR (2013) Sustainable production of bio-based chemicals by extremophiles. Curr Biotechnol 2:360–379. doi:10.2174/18722083113076660028

Enfors SO, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, Krüger E, Schweder T, Hamer G, O'Beirne D, Noisommit-Rizzi N, Reuss M, Boone L, Hewitt C, McFarlane C, Nienow A, Kovacs T, Trägårdh C, Fuchs L, Revstedt J, Friberg PC, Hjertager B, Blomsten G, Skogman H, Hjort S, Hoeks F, Lin HY, Neubauer P, van der Lans R, Luyben K, Vrabel P, Manelius AL (2001) Physiological responses to mixing in large scale bioreactors. J Biotechnol 85:175–185. doi:10.1016/S0168-1656(00)00365-5

Fahnert B, Lilie H, Neubauer P (2004) Inclusion bodies: formation and utilisation. Adv Biochem Eng Biotechnol 89:93–142. doi:10.1007/b93995

Feijoo-Siota L, Blasco L, Rodríguez-Rama JL, Barros-Velázquez J, Miguel T, Sánchez-Pérez A, Villa TG (2014) Recent patents on microbial proteases for the dairy industry. Recent Adv DNA Gene Seq 8:44–55. doi:10.2174/2352092208666141013231720

Fisher TE, Oberhauser AF, Carrion-Vazquez M, Marszalek PE, Fernandez JM (1999) The study of protein mechanics with the atomic force microscope. Trends Biochem Sci 24:379–384. doi:10.1016/S0968-0004(99)01453-X

Forman JR, Clarke J (2007) Mechanical unfolding of proteins: insights into biology, structure and folding. Curr Opin Struct Biol 17:58–66. doi:10.1016/j.sbi.2007.01.006

Fornwald JA, Lu Q, Boyce FM, Ames RS (2016) Gene expression in mammalian cells using BacMam, a modified Baculovirus system, Baculovirus and insect cell expression protocols. In: Murhammer DW (ed). Methods Mol Biol 1350:95–116. doi:10.1007/978-1-4939-3043-2_5

Fulwyler MJ (1965) Electronic separation of biological cells by volume. Science 150:910–911. doi:10.1126/science.150.3698.910

Gandhi P, Page M, Chen Z, Bush-Pelc L, Di Cera E (2009) Mechanism of the anticoagulant activity of thrombin mutant W215A/E217A. J Biol Chem 284:24098–24105. doi:10.1074/jbc.M109.025403

Garcia R, Perezb R (2002) Dynamic atomic force microscopy methods. Surf Sci Rep 47:197–301. doi:10.1016/S0167-5729(02)00077-8

Garcia-Carreon FL, Del Toro MAN (1997) Classification of proteases without tears. Biochem Educ 25:161–167. doi:10.1016/S0307-4412(97)00005-8

Gebauer M, Skerra A (2009) Engineered protein scaffolds as next-generation antibody therapeutics. Curr Opin Chem Biol 13:245–255. doi:10.1016/j.cbpa.2009.04.627

Genest O, Reidy M, Street TO, Hoskins JR, Camberg JL, Agard DA, Masison DC, Wickner S (2013) Uncovering a region of heat shock protein 90 important for client binding in *E. coli* and chaperone function in yeast. Mol Cell 49:464–473. doi:10.1016/j.molcel.2012.11.017

Genevaux P, Georgopoulos C, Kelley WL (2007) The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. Mol Microbiol 66:840–857. doi:10.1111/j.1365-2958.2007.05961.x

Georgiou G, Valax P (1996) Expression of correctly folded proteins in *Escherichia coli*. Curr Opin Biotechnol 7:190–197. doi:10.1016/S0958-1669(96)80012-7

Ginn C, Khalili H, Lever R, Brocchini S (2014) PEGylation and its impact on the design of new protein-based medicines. Future Med Chem 6:1829–1846. doi:10.4155/fmc.14.125

Glover JR, Lindquist S (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. Cell 94:73–82. doi:10.1016/S0092-8674(00)81223-4

Goldsmith M, Tawfik DS (2012) Directed enzyme evolution: beyond the low-hanging fruit. Curr Opin Struct Biol 22:406–412. doi:10.1016/j.sbi.2012.03.010

Grubmuller H, Heymann B, Tavan P (1996) Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. Science 271:997–999. doi:10.1126/science.271.5251.997

Guerrero JL, Daugherty PS, O'Malley MA (2016) Emerging technologies for protease engineering: new tools to clear out disease. Biotechnol Bioeng. doi:10.1002/bit.26066

Gupta P, Aggarwal N, Batra P, Mishra S, Chaudhuri TK (2006) Co-expression of chaperonin GroEL/GroES enhances in vivo folding of yeast mitochondrial aconitase and alters the growth characteristics of *Escherichia coli*. Int J Biochem Cell Biol 38:1975–1985. doi:10.1016/j.biocel.2006.05.013

Gupta R, Beg QK, Lorenz P (2002) Bacterial alkaline proteases: molecular approaches and industrial applications. Appl Microbiol Biotechnol 59:15–32. doi:10.1007/s00253-002-0975-y

Müller KM, Stebel SC, Knall S, Zipf G, Bernauer HS, Arndt KM (2005) Nucleotide exchange and excision technology (NExT) DNA shuffling: a robust method for DNA fragmentation and directed evolution. Nucleic Acids Res 33:e117. doi:10.1093/nar/gni116

Nagata Y, Chu K H (2003) Optimization of a fermentation medium using neural networks and genetic algorithms. Biotechnol Lett 25: 1837–1842. doi: 10.1023/A:1026225526558.

Nettels D, Gopich IV, Hoffmann A, Schuler B (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. Proc Natl Acad Sci U S A 104:2655–2660. doi:10.1073/pnas.0611093104

Neubauer P et al (2013) Consistent development of bioprocesses from microliter cultures to the industrial scale. Eng Life Sci 13:224–238. doi:10.1002/elsc.201200021

Neupane K, Solanki A, Sosova I, Belov M, Woodside MT (2014) Diverse metastable structures formed by small oligomers of a-synuclein probed by force spectroscopy. PLoS One 9:e86495. doi:10.1371/journal.pone.0086495

Neurath H, Walsh KA (1976) Role of the proteolytic enzymes in biological regulation (a review). Proc Natl Acad Sci U S A 73:3825–3832

Novotny L, Bian RX, Xie XS (1997) Theory of nanometric optical tweezers. Phys Rev Lett 79:645–654. doi:10.1103/PhysRevLett.79.645

O'Beirne D, Hamer G (2000) Oxygen availability and the growth of *Escherichia coli* W-3110: a problem exacerbated by scale up. Biprocess Eng 23:487–494. doi:10.1007/s004499900185

Ogunbiyi L, Riedhammer TM, Smith FX (1986) Method for enzymatic cleaning and disinfecting contact lenses. US Patent 4614549

Olama ZA, el-Sabaeny AH (1993) Lipase production by *Aspergillus niger* under various growth conditions using solid state fermentation. Microbiologia 9:134–141

Ostermeier M, Shim JH, Benkovic SJ (1999) A combinatorial approach to hybrid enzymes independent of DNA homology. Nat Biotechnol 17:1205–1209. doi:10.1038/70754

Ozturk SS (1996) Engineering challenges in high density cell culture systems. Cytotechnology 22:3–16. doi:10.1007/BF00353919

Palomares LA, Ramírez OT (1996) The effect of dissolved oxygen tension and the utility of oxygen uptake rate in insect cell culture. Cytotechnology 22:225–237. doi:10.1007/BF00353943

Palomares LA, Ramírez OT (2003) Bioreactor scaleup. In: Spier RE (ed) The encyclopedia of cell technology. Wiley, New York, pp 174–183. doi:10.1002/0471250570.spi023

Pantoliano MW, Whitlow M, Wood JF, Rollence ML, Finzel BC, Gilliland GL, Poulos TL, Bryan PN (1988) The engineering of binding affinity at metal ion binding sites for the stabilization of proteins: subtilisin as a test case. Biochemist 27:8311–8317. doi:10.1021/bi00422a004

Pavlou AK, Reichert JM (2004) Recombinant protein therapeutics-success rates, market trends and values to 2010. Nat Biotechnol 22:1513–1519. doi:10.1038/nbt1204-1513

Persson E, Kjalke M, Olsen O (2001) Rational design of coagulation factor VIIa variants with substantially increased intrinsic activity. Proc Natl Acad Sci U S A 98:13583–13588. doi:10.1073/pnas.241339498

Peterman EJ, Sosa H, Moerner WE (2004) Single-molecule fluorescence spectroscopy and microscopy of biomolecular motors. Annu Rev Phys Chem 55:79–96. doi:10.1146/annurev.physchem.55.091602.094340

Pisal DS, Kosloski MP, Balu-Iyer SV (2010) Delivery of therapeutic proteins. J Pharm Sci 99:2557–2575. doi:10.1002/jps.22054

Piston DW, Kremers GJ (2007) Fluorescent protein FRET: the good, the bad and the ugly. Trends Biochem Sci 32:407–414. doi:10.1016/j.tibs.2007.08.003

Porter JL, Rusli RA, Ollis DL (2016) Directed evolution of enzymes for industrial biocatalysis. Chembiochem 17:197–203. doi:10.1002/cbic.201500280

Radner S, Celie PH, Fuchs K, Sieghart W, Sixma TK, Stornaiuolo M (2012) Transient transfection coupled to baculovirus infection for rapid protein expression screening in insect cells. J Struct Biol 179:46–55. doi:10.1016/j.jsb.2012.04.013

Rai M, Padh H (2001) Expression systems for production of heterologous proteins. Curr Sci 80:1121–1128

Rajan RS, Tsumoto K, Tokunaga M, Tokunaga H, Kita Y, Arakawa T (2011) Chemical and pharmacological chaperones: application for recombinant protein production and protein folding diseases. Curr Med Chem 18:1–15. doi:10.2174/092986711793979698

Rao MB, Tanskale AM, Ghatger MS, Deshpande VV (1998) Molecular and biotechnological aspects of microbial proteases. Microbiol Mol Biol Rev 62:597–635

Ravikumar A, Arrieta A, Liu CC (2014) An orthogonal DNA replication system in yeast. Nat Chem Biol 10:175–177. doi:10.1038/nchembio.1439

Reif MM, Oostenbrink C (2014) Molecular dynamics simulation of configurational ensembles compatible with experimental FRET efficiency data through a restraint on instantaneous FRET efficiencies. J Comput Chem 35:2319–2332. doi:10.1002/jcc.23756

Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE (2013) Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. Science 339:1080–1083. doi:10.1126/science.1233066

Sakmann B, Neher E (1984) Patch clamp techniques for studying ionic channels in excitable membranes. Annu Rev Physiol 46:455–472. doi:10.1146/annurev.ph.46.030184.002323

Sandhya C, Sumantha A, Szakacs G, Pandey A (2005) Comparative evaluation of neutral protease production by *Aspergillus oryzae* in submerged and solid-state fermentation. Process Biochem 40:2689–2694. doi:10.1016/j.procbio.2004.12.001

Schlieker C, Bukau B, Mogk A (2002) Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: implications for their applicability in biotechnology. J Biotechnol 96:13–21. doi:10.1016/S0168-1656(02)00033-0

Schmidt FR (2005) Optimization and scale up of industrial fermentation processes. Appl Microbiol Biotechnol 68:425–435. doi:10.1007/s00253-005-0003-0

Shankar S, More SV, Laxman RS (2010) Recovery of silver from waste x-ray film by alkaline protease from *Conidiobolus Coronatus*. Kathmandu Univ J Sci Eng Technol 6:60–69. doi:10.3126/kuset.v6i1.3311

Shuler ML, Kargi F (2002) Bioprocess engineering. Basic concepts, 2nd edn. Prentice Hall, Upper Saddle River

Skerra A (2000) Engineered protein scaffolds for molecular recognition. J Mol Recognit 13:167–187. doi:10.1002/1099-1352(200007/08)13:4<167::AID-JMR502>3.0.CO;2-9

Smith SB, Cui Y, Bustamante C (2003) Optical-trap force transducer that operates by direct measurement of light momentum. Methods Enzymol 361:134–162

Snow CD, Sorin EJ, Rhee YM, Pande VS (2005) How well can simulation predict protein folding kinetics and thermodynamics? Annu Rev Biophys Biomol Struct 34:43–69. doi:10.1146/annurev.biophys.34.040204.144447

Srivastava A, Arya P, Goel S, Kundu B, Mishra P, Fnu A (2015) Gelsolin amyloidogenesis is effectively modulated by curcumin and emetine conjugated PLGA nanoparticles. PLoS One 10:e0127011. doi:10.1371/journal.pone.0127011

Stangner T, Wagner C, Singer D, Angioletti-Uberti S, Gutsche C, Dzubiella J, Hoffmann R, Kremer F (2013) Determining the specificity of monoclonal antibody HPT-101 to tau-peptides with optical tweezers. ACS Nano 7:11388–11396. doi:10.1021/nn405303u

Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389–391. doi:10.1038/370389a0

Strausberg SL, Alexander PA, Gallagher DT, Gilliland GL, Barnett BL, Bryan PN (1995) Directed evolution of a subtilisin with calcium-independent stability. Nat Biotechnol 13:669–673. doi:10.1038/nbt0795-669

Structural Genomics Consortium, China Structural Genomics Consortium, Northeast Structural Genomics Consortium, Gräslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, Prilusky J, Sussman JL, Stevens RC, Lesley SA, Wilson IA, Joachimiak A, Collart F, Dementieva I, Donnelly MI, Eschenfeldt WH, Kim Y, Stols L, Wu R, Zhou M, Burley SK, Emtage JS, Sauder JM, Thompson D, Bain K, Luz J, Gheyi T, Zhang F, Atwell S, Almo SC, Bonanno JB, Fiser A, Swaminathan

S, Studier FW, Chance MR, Sali A, Acton TB, Xiao R, Zhao L, Ma LC, Hunt JF, Tong L, Cunningham K, Inouye M, Anderson S, Janjua H, Shastry R, Ho CK, Wang D, Wang H, Jiang M, Montelione GT, Stuart DI, Owens RJ, Daenke S, Schütz A, Heinemann U, Yokoyama S, Büssow K, Gunsalus KC (2008) Protein production and purification. Nat Methods 5:135–146. doi:10.1038/nmeth.f.202

Tezuka-Kawakami T, Gell C, Brockwell DJ, Radford SE, Smith DA (2006) Urea-induced unfolding of the immunity protein Im9 monitored by spFRET. Biophys J 91:L42–L44. doi:10.1529/biophysj.106.088344

Theron LW, Divol B (2014) Microbial aspartic proteases: current and potential applications in industry. Appl Microbiol Biotechnol 98:8853–8868. doi:10.1007/s00253-014-6035-6

Thomas JG, Baneyx F (1996) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells over expressing heat-shock proteins. J Biol Chem 271:11141–11147. doi:10.1074/jbc.271.19.11141

Thomas JG, Baneyx F (1997) Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*. Protein Expr Purif 11:289–296. doi:10.1006/prep.1997.0796

Thomas T, Gilbert J, Meyer F (2012) Metagenomics—a guide from sampling to data analysis. Microb Inform Exp 2:3. doi:10.1186/2042-5783-2-3

Tian H, Fürstenberg A, Huber T (2017) Labeling and single-molecule methods to monitor G protein-coupled receptor dynamics. Chem Rev 117(1):186–245

Tripathi NK, Sathyaseelan K, Jana AM, Rao PVL (2009) High yield production of heterologous proteins with *Escherichia coli*. Def Sci J 59:137–146

van Beilen JB, Li Z (2002) Enzyme technology: an overview. Curr Opin Biotechnol 13:338–344. doi:10.1016/S0958-1669(02)00334-8

Vázquez-Rey M, Lang DA (2011) Aggregates in monoclonal antibody manufacturing processes. Biotechnol Bioeng 108:1494–1508. doi:10.1002/bit.23155

Ventura S, Villaverde A (2006) Protein quality in bacterial inclusion bodies. Trends Biotechnol 24:179–185. doi:10.1016/j.tibtech.2006.02.007

Villaverde A, Carrio MM (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnol Lett 25:1385–1395. doi:10.1023/A:1025024104862

Vishalakshi N, Lingappa K, Amena S, Prabhakar M, Dayanand A (2009) Production of alkaline protease from *Streptomyces gulbargensis* and its application in removal of blood stain. Ind. J Biotechnol 8:280–285

Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM (1998) Force and velocity measured for single molecules of RNA polymerase. Science 282:902–907. doi:10.1126/science.282.5390.902

Wang Z, Moult J (2001) SNPs, protein structure, and disease. Hum Mutat 17:263–270. doi:10.1002/humu.22

Watkins LP, Chang H, Yang H (2006) Quantitative single-molecule conformational distributions: a case study with poly-(L-proline). J Phys Chem A 110:5191–5203. doi:10.1021/jp055886d

Weckwerth W (2011) Green systems biology—from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. J Proteome 75:284–305. doi:10.1016/j.jprot.2011.07.010

Xia J, Wang G, Lin J, Wang Y, Chu J, Zhuang Y, Zhang S (2015) Advances and practices of bioprocess scale-up. Adv Biochem Eng Biotechnol 152:137–151. doi:10.1007/10_2014_293.

Yasuhara A, Ogawa A, Tanaka T, Sakiyama T, Nakanishi K (1994) Production of neutral protease from *Aspergillus oryzae* by a novel cultivation method on a microporous membrane. Biotechnol Tech 8:249–254. doi:10.1007/BF00155416

Yegneswaran PK, Thompson BG, Gray MR (1991) Effect of dissolved oxygen control on growth and antibiotic production by *Strepmomyces clavuligerus* fermentations. Biotechnol Prog 7:246–250. doi:10.1021/bp00009a008

Young JC, Agashe R V, Siegers K, Hartl FU (2004) Pathways of chaperone-mediated protein folding in the cytosol. Nat Rev Mol Cell Biol 5:781–791. doi:10.1038/nrm1492

Young JC, Barral JM, Ulrich Hartl F (2003) More than folding: localized functions of cytosolic chaperones. Trends Biochem Sci 28:541–547. doi:10.1016/j.tibs.2003.08.009

Yu H, Dee DR, Liu X, Brigley AM, Sosova I, Woodside MT (2015) Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape. Proc Natl Acad Sci U S A 112:8308–8313. doi:10.1073/pnas.1419197112

Yuan C, Chen A, Kolb P, Moy T V (2000) Energy landscape of streptavidin-biotin complexes measured by atomic force microscopy. Biochemistry 39:10219–10223. doi:10.1021/bi992715o

Zhao H, Zha W (2004) Evolutionary methods for protein engineering. Marcel Dekker, Inc., New York, pp 353–373

Zhao HY, Wu LY, Liu G, Feng H (2016) Single-site substitutions improve cold activity and increase thermostability of the dehairing alkaline protease (DHAP). Biosci Biotechnol Biochem 80:2480–2485. doi:10.1080/09168451.2016.1230005

Zolkiewski M (1999) ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from *Escherichia coli*. J Biol Chem 274:28083–28086. doi:10.1074/jbc.274.40.28083