7 Engineering of Therapeutic Proteins

Fei Wen, Sheryl B. Rubin-Pitel, and Huimin Zhao

CONTENTS

Protein Therapeutics versus Small Molecule Drugs .............................................. 154
Sources of Protein Therapeutics ................................................................... 155
Targets of Protein Therapeutics and Modes of Action.................................. 155
Engineering Effective Protein Therapeutics .......................................................... 156
Challenges in Pharmaceutical Translation of New Therapeutic Proteins.... 156
Strategies for Designing Effective Protein Therapeutics .............................. 156
Strategies for Improving Pharmacokinetics...................................................... 157
Strategies for Reducing Immunogenicity ......................................................... 158
Genetic Engineering ...................................................................................... 159
Examples of Protein Therapeutics ......................................................................... 160
Monoclonal Antibody Therapeutics.............................................................. 161
Enzyme Therapeutics Acting on Extracellular Targets ...................................... 161
Protein Therapeutics as Replacements for Defective or Deficient Proteins.... 162
Protein Hormones ....................................................................................... 162
Coagulation Factors ..................................................................................... 163
Enzyme Replacement Therapy ......................................................................... 163
Cytokines and Their Receptors as Protein Therapeutics .............................. 164
Colony Stimulating Factors ........................................................................... 164
Interleukins and Their Receptors ...................................................................... 164
Interferons ....................................................................................................... 165
New Classes of Therapeutic Proteins under Development .................................... 167
Major Histocompatibility Complex (MHC) Proteins ................................... 167
Phage Display ................................................................................................. 167
Cell Surface Display ....................................................................................... 167
T Cell Receptors (TCRs) ............................................................................... 168
Protein Fusions and Rational Design .............................................................. 169
Phage Display ................................................................................................. 170
Yeast Surface Display ..................................................................................... 170
TCR-Like Antibodies ....................................................................................... 170
Conclusion and Future Prospects ...................................................................... 171
References .............................................................................................................. 171
We live in a pharmaceutical age with access to more medicines than any other time in human history. Among these, there are two major classes of drugs in the market: small molecule drugs and protein therapeutics. Although the former is currently the principal therapeutic agent, the impact of protein therapeutics is increasing thanks to advances in recombinant DNA technology and improved understanding of disease pathology. Currently, the U.S. Food and Drug Administration (FDA) has approved over 150 different protein drugs including monoclonal antibodies (mAbs), insulin, erythropoietin, interferons, and somatropin (human growth hormone) (Wishart et al. 2008). The protein therapeutics market is continuously growing and has more than doubled in the past few years, jumping from $25 billion in 2001 to $57 billion in 2006 with strong sales in insulin, erythropoietin, and interferon segments (KaloramaInformation 2006; RNCOS 2007). In addition to small molecule drugs and protein therapeutics, there are also other approaches being researched, including gene therapy, RNA interference (RNAi), stem cells, and nanotechnology-based solutions. These technologies hold great promise as future pharmaceuticals, but they are still a long way from routine use. This chapter will mainly focus on therapeutic proteins that are either in clinical use or under development. After a brief overview of the molecular basis of protein therapeutics, we will discuss some recent advances in protein engineering and design technologies, with an emphasis on diversity-oriented methods and their applications to protein therapeutics.

PROTEIN THERAPEUTICS VERSUS SMALL MOLECULE DRUGS

Small molecule drugs have several advantages, including oral bioavailability, ability to reach intracellular targets, ease of manufacturing, and generally a long shelf life. These characteristics make them favorable over protein drugs in pharmaceutical industry (see the section titled “Challenges in Pharmaceutical Translation of New Therapeutic Proteins,” in this chapter). However, small molecule drugs, which typically have a molecular weight less than 1000 Daltons, have limited surface area available to contact a target protein. Furthermore, forming a favorable interaction requires the presence of a deep hydrophobic pocket in the target protein, limiting the number of potential druggable targets (Hopkins and Groom 2002; Hopkins and Groom 2003). In contrast, protein drugs are usually large in size and do not have this limitation, making them indispensable therapeutic tools for human disease treatment. By using protein therapeutics, some debilitating diseases that were previously untreatable, such as chronic renal failure, dwarfism, and infertility, are now successfully managed (Johnson-Leger et al. 2006).

Protein therapeutics have higher binding selectivity and specificity compared to small molecule drugs; therefore they can target specific steps in disease pathology. For example, before the advent of protein therapeutics, drugs used to suppress the immune system in chronic inflammatory disorders were limited to small molecule drugs, such as corticosteroids and cyclosporine A. These drugs act broadly and inhibit both protective and harmful immune responses indiscriminately, thus having serious side effects. In contrast, mAbs such as Infliximab (Remicade®, Centocor Inc) are considered immune-modulating. Infliximab targets tumor necrosis factor-α, a key proinflammatory cytokine in the pathogenesis of chronic immune disorders,
and leaves the protective immune response intact (Rutgeerts et al. 2006). In the past
decades, the development of new protein therapeutics has revised the treatment para-
digm of certain diseases (Flamant and Bourreille 2007; Gergely and Fekete 2007),
and they are gradually replacing or supplementing small molecule drug therapies
(Johnson-Leger et al. 2006; Eng 2007; Flamant and Bourreille 2007; Gergely and
Fekete 2007).

**SOURCES OF PROTEIN THERAPEUTICS**

The human body has evolved an elegant immune system that helps combat and con-
trol diseases. Insufficient, deficient, or improper action of any molecular component
of the immune system results in disorders to various extents. Therefore, extrinsic
modulation of the immune system using natural human immune regulators repre-
sents an appealing strategy to cure diseases, and the human genome, completely
sequenced in 2003, provides a huge source for drug target mining. In fact, before the
introduction of recombinant DNA technology, therapeutic proteins such as growth
hormone and follicle stimulating hormone were isolated directly from the human
body. With the advance of recombinant DNA technology, therapeutic proteins in the
market now include recombinant antibodies, hormones, cytokines, interferons, and
enzymes of human origin produced industrially in bacterial, yeast, or mammalian
expression systems (Johnson-Leger et al. 2006).

In addition to human immunoregulatory proteins, many viruses are also “experts”
at manipulating the human immune system to facilitate their propagation in the host.
Viruses evade or subvert host immune detection and destruction by encoding and
expressing a diverse array of immunomodulatory proteins, which target pathways of
antibody response, cytokine-mediated signaling, and major histocompatibility com-
plex (MHC)—restricted antigen presentation (Tortorella et al. 2000; Alcami 2003).
After eons of coevolution, these virus-engineered immunomodulatory proteins have
exquisite potency and specificity unrivaled by commercial pharmaceuticals, provid-
ing a powerful platform of protein therapeutics development (Lucas and McFadden
2004). Therefore, similar to the concept of human genome mining described previ-
ously, human “virome” mining has been proposed to uncover more drug candidates
(Anderson et al. 2003; DeFilippis et al. 2003).

**TARGETS OF PROTEIN THERAPEUTICS AND MODES OF ACTION**

Any molecule that has implications in the pathogenesis of a disease is a potential target
for protein therapeutics. In contrast to small molecule drugs that are able to diffuse
across cell membranes, protein therapeutics typically cannot traverse this cellular barrier
due to their large size. Therefore, they almost exclusively target cell surface receptors or
extracellular molecules. In recent years, researchers have also explored the possibility of
directing protein therapeutics to intracellular targets (Stocks 2004; Bernal et al. 2007).

Broadly speaking, protein therapeutics have three different modes of action based
on the pathology of a disease. First, if the disease is caused by unwanted extracellular
molecules such as cell metabolites or cell lysate, enzyme therapeutics can degrade
these targets. Second, if the disease is caused by a deficiency in certain proteins,
such as enzymes, protein therapeutics can be used to replace them and restore an individual’s health. Third, if the disease involves improper immune responses or dysregulated signaling pathways, such as chronic inflammatory diseases, autoimmune diseases, infectious diseases, and cancers, protein therapeutics act as inhibitors or activators of cell surface receptors. Among these three categories, the last one has attracted the most attention of researchers and represents an active area of research (see the section titled “Examples of Therapeutic Proteins,” in this chapter).

ENGINEERING EFFECTIVE PROTEIN THERAPEUTICS

Protein therapeutics clearly have indisputable importance among modern pharmaceuticals. However, they remain at an early stage of development and application, and substantial improvement must be made in almost all aspects, including drug target identification, protein engineering and design, protein expression and purification, drug delivery, and marketing. Here we will focus on recent advances in protein engineering and design technologies.

CHALLENGES IN PHARMACEUTICAL TRANSLATION OF NEW THERAPEUTIC PROTEINS

Although the human genome represents a rich source of candidate proteins for therapeutics, these proteins were not evolved for therapeutic purposes, and thus do not have optimal affinity, specificity, activity, and/or stability for disease treatment. Protein instability and immunogenicity are among the key challenges affecting the success of protein therapeutics. Proteins often have limited physical and chemical stability, which has several implications: (1) They have short half-life in the body, which in turn leads to limited efficacy and frequent dosage; (2) they are difficult to produce and have short shelf life, both of which are responsible for high cost of pharmaceutical commercialization; (3) they need to be administered through injection because they are quickly digested in the intestines if taken orally, which affects patient compliance and therapeutic outcomes (Kefalides 1998). Therefore, development of protein therapeutics with improved stability, efficacy, pharmacokinetics, pharmacodynamics, and expression productivity is required.

Safety is the priority criterion of all drugs. Immunogenicity is a unique issue associated with protein therapeutics. The human immune system responds to pathogens through recognition of their proteins or processed protein products. Similarly, patients who receive protein therapeutics can potentially develop immune responses against the protein drug, producing antidrug antibodies (Barbosa and Celis 2007; De Groot and Scott 2007). Such immune responses can reduce the efficacy of protein drugs, and in rare cases, they can lead to life-threatening situations (Barbosa and Celis 2007; De Groot and Scott 2007). Therefore, an immunogenicity assessment is essential to protein therapeutics development.

STRATEGIES FOR DESIGNING EFFECTIVE PROTEIN THERAPEUTICS

To address the issues in protein therapeutics development described previously, various protein engineering and design strategies have been developed, including protein
post-translational modification, protein fusions, and genetic engineering, as illustrated in Figure 7.1.

**Strategies for Improving Pharmacokinetics**

The efficacy of a therapeutic protein in the human body can be improved by a number of strategies, including fusions, glycosylation, and chemical modification. Recombinant DNA technology enables the fusion of a protein of interest to an endogenous human protein (for example, human serum albumin, an antibody Fc fragment, and transferrin), thereby increasing the effective size of the protein and reducing clearance in the kidneys, which occurs for proteins below a molecular weight of approximately 70,000 Daltons (Caliceti and Veronese 2003; Beals and Shanafelt 2006). Fusing a protein to an antibody Fc fragment affords a second benefit: This takes advantage of the interaction between Fc and the FcRn receptor, which protects IgG antibodies in the body by allowing them to be released into the plasma rather than degraded in the lysosome, a natural “recycling” system (Lobo et al. 2004).

*Glycosylation*, the term for decoration of a protein’s surface with carbohydrates, also increases protein size, thus reducing renal clearance. In addition, glycosylation can further help enhance protein solubility, stabilize against damage from heat and free radicals, and shield a protein from proteolysis and immune surveillance, ultimately resulting in enhanced serum half-life (Sinclair and Elliott 2005). Targeted introduction of glycosylation motifs into the sequence of a therapeutic protein is termed *glycoengineering*. Production of recombinant protein by mammalian cell culture, particularly Chinese hamster ovary (CHO) cells, is used industrially to synthesize glycosylated therapeutic proteins. The use of CHO cell culture for protein production does have drawbacks, namely significantly higher cost versus culture of

---

**FIGURE 7.1** Strategies for designing effective protein therapeutics. StEP = staggered extension process, ITCHY = incremental truncation for the creation of hybrid enzymes, SHIPREC = sequence homology-independent protein recombination (see Chapter 4 for a detailed description of each method).
bacteria or yeast. It is also possible that CHO cell cultures can harbor viral or prion contamination, and the glycosylation by CHO cells can be heterogeneous, leading to therapeutic proteins with a range of efficacies (Sethuraman and Stadheim 2006).

Finally, chemical modification by PEGylation, the conjugation of polyethylene glycol (PEG), is a common strategy to enhance the serum half-life of protein therapeutics (Beals and Shanafelt 2006). The increased size of the protein-PEG conjugate reduces clearance from the kidneys, and the bulky PEG molecule also protects therapeutic proteins from degradation by proteases via steric hindrance (Veronese and Pasut 2005). However, PEGylation also suffers from several limitations. Low molecular weight proteins are especially susceptible to partial or complete inactivation after conjugation of a PEG molecule (Shechter et al. 2008). Further, interference of protein-protein binding is beneficial vis-à-vis proteases, but reduces the effectiveness of antibody-based protein therapeutics or those acting through a receptor (Kubetzko et al. 2005; Shechter et al. 2008). This has led to the development of “reversible PEGylation,” where the conjugated PEG molecule can be considered a prodrug, and undergoes spontaneous hydrolysis under physiological conditions to release the active therapeutic protein (Shechter et al. 2008).

Strategies for Reducing Immunogenicity

In general, reduction of immunogenicity involves altering protein therapeutics such that they can avoid immune surveillance. This includes avoidance of antibodies, binding to antigen presenting cell (APC) surface receptors leading to receptor-mediated endocytosis, subsequent proteolysis to peptide fragments that bind to MHC class II molecules, and, finally, avoidance of binding by B and T cell receptors (Chirino et al. 2004). Many strategies are available to reduce immunogenicity and are similar to techniques used to improve pharmacokinetics. PEGylation, which possesses other useful properties as discussed previously, is nontoxic and reduces immunogenicity and antigenicity (Caliceti and Veronese 2003). Specifically, the PEG molecule shields immunoreactive sites on recombinant proteins from recognition by antibodies or surface receptors (Caliceti and Veronese 2003). PEG also enhances solubility, which prevents the accumulation of highly immunogenic protein aggregates (Chirino et al. 2004). Conjugated PEG also deters proteolysis, which may help PEGylated therapeutics to avoid cleavage into peptides capable of display on MHC class II molecules.

Glycosylation, which helps to enhance the serum half-life of therapeutic proteins, is also thought to interfere with antibody binding (De Groot and Scott 2007). Ideally, glycosylation should be of a human pattern for greatest effectiveness (Brooks 2006). This has led to considerable interest in humanizing the glycosylation pathways of the organisms currently used to express recombinant proteins, and fully humanized, sialylated glycoproteins can be produced from the yeast *Pichia pastoris* (Hamilton et al. 2003). Changes to the primary sequence of a therapeutic protein can also help reduce immunogenicity. The humanization of murine and mouse-human chimeric antibodies (i.e., the removal of as much nonhuman content from the constant and variable regions as possible) helps to reduce the formation of human antimouse antibodies (De Groot and Scott 2007). In addition, identifying and eliminating antibody and T-cell epitopes and class II MHC agretopes are strategies currently employed to
reduce the immunogenicity of the next generation of protein therapeutics (Chirino et al. 2004).

**Genetic Engineering**

Genetic engineering strategies, consisting of three broad categories—rational design, directed evolution, and semirational design—have long been valuable tools in engineering proteins with altered physical and chemical properties and/or creating novel functions. In the context of protein therapeutics, these well-established methods are facing new challenges due to consideration of additional engineering parameters including pharmacokinetics, pharmacodynamics, and immunogenicity as described in the section in this chapter titled “Challenges in Pharmaceutical Translation of New Therapeutic Proteins.” Numerous rational and semirational design strategies have been developed (detailed in other chapters in this volume) and applied to the engineering of protein therapeutics. Rational or computation design methods have been applied to improve stability and solubility, or to predict and to reduce immunogenicity of protein therapeutics (reviewed in Marshall et al. 2003; Rosenberg and Goldblum 2006; De Groot and Moise 2007). The primary drawback of rational design is the requirement for knowledge of protein structure, mechanisms, and protein structure-function relationships to a certain extent. In contrast, directed evolution does not have this limitation because it creates molecular diversity at the DNA level in a stochastic manner (see Chapter 4). This also leads to the key challenge of directed evolution: how to find the variant with desired property in a library of up to a billion variants. Therefore, high-throughput selection or screening methods are highly desirable for directed evolution. A variety of library selection and screening methods have been developed for different applications (reviewed in Arnold 2003). For each directed evolution experiment, the selection or screening method must be prudently chosen or developed, because the first principle of directed evolution is “you get what you select (screen) for.”

Among the library selection and screening approaches, display technologies have been increasingly used in therapeutic protein engineering, and have proved to be especially powerful for engineering protein drugs for improved affinity and specificity. The shared principle of different display technologies is to create a physical linkage between the genotype and the protein displayed on the platform, so that a library of target protein variants is directly accessible to binding analysis and thus selectable and recoverable for further engineering. By using surface display, *in vitro* affinity maturation of an antibody yielded variants with the highest affinity reported (femtomolar range, which is orders of magnitude beyond natural antibodies) (Boder et al. 2000), and new classes of therapeutic proteins are being developed (see the section titled “New Classes of Therapeutic Proteins under Development,” in this chapter). Over the past decades, a number of display platforms have been developed, including phage display, cell surface display, and cell-free display (see Chapters 1–3). These different platforms have advantages and disadvantages that make them more conducive to certain protein engineering applications.

Phage display was the earliest developed platform and has since been utilized most often for protein engineering (Sergeeva et al. 2006). Recent advances have enabled selection of phage libraries in more complex biological systems, such as cultured
cells and in vivo (Sergeeva et al. 2006). Although phage display has been successfully used for engineering of peptides, antibodies, and for epitope mapping, it has achieved limited success with more complex human membrane proteins such as MHC and T cell receptors (TCRs) (see the section titled “New Classes of Therapeutic Proteins under Development,” in this chapter). This is because the bacterial host required for phage propagation has limited ability in terms of protein folding and post-translational modifications that are important for mammalian protein functions.

A cell surface display library is usually generated by transforming cells with DNA variants and screening for mutants with a desired phenotype by fluorescence activated cell sorting (FACS). FACS enables high-throughput enrichment of positive clones in a quantitative manner, but is not applicable for phage display libraries due to the small size of phage particles (Georgiou et al. 1997; Boder and Wittrup 1998). Several different cell types have been explored for their ability to display protein libraries, including bacteria, yeast, insect, and mammalian cells. Among these platforms, yeast display has attracted the most attention. Yeast display has the advantage of possessing post-translational processing pathways, which enable folding and glycosylation of complex human proteins (Kondo and Ueda 2004). Starting with the same library, yeast display was shown to sample the immune antibody repertoire considerably more fully than phage display, selecting twice as many novel antibodies as phage display (Bowley et al. 2007). Studies have also shown that the surface display level of a protein on yeast cell surface is strongly correlated with both thermal stability and soluble expression level (Shusta et al. 1999).

Cell-free display (also known as in vitro display) represents an emerging technology that has proven useful for discovery and engineering of therapeutic proteins with high affinity (FitzGerald 2000; Rothe et al. 2006). For example, ribosome (polysome) display, which was the first cell-free system developed for complete in vitro protein engineering, has been shown to generate antibodies with higher affinities (picomolar range) than those obtained from a phage display library (nanomolar range) (Groves et al. 2006). The biggest advantage of cell-free display methods is that the transcription and/or translation steps completely take place in vitro, abolishing the need of introducing DNA into host cells, which often limits the library size accessible to other display approaches. Library sizes created by in vitro display platforms are usually several orders of magnitude higher than that obtained with other display methods (up to $10^{14}$–$10^{15}$) (FitzGerald 2000). In addition, the cell-free feature of in vitro display methods might make them more amenable to automation, potentially allowing ultra high-throughput identification of new drug targets on a genomic level (FitzGerald 2000).

EXAMPLES OF PROTEIN THERAPEUTICS

The previously described protein engineering and design strategies have been applied to engineer a wide variety of protein therapeutics for enhanced activity, stability, affinity, specificity, pharmacokinetics, pharmacodynamics, reduced immunogenicity, and improved productivity.
MONOCLONAL ANTIBODY THERAPEUTICS

Extensive review articles and book volumes have been devoted to the subject of monoclonal antibodies and antibody engineering; therefore, a detailed discussion will not be provided here. However, it is worth mentioning that their exquisite specificity, ease of engineering fragments by display technologies, and chimerization or humanization to enhance stability have propelled antibodies to become a swiftly expanding class of therapeutics for treating a wide variety of human conditions [46]. Future directions of research include the development of multiantibody cocktails for synergistic effects (Logtenberg 2007) and the conjugation of antibodies to immunotoxic drugs to enhance tumor killing (Zafir-Lavie et al. 2007).

ENZYME THERAPEUTICS ACTING ON EXTRACELLULAR TARGETS

In addition to binding proteins targeting soluble molecules and membrane-bound receptors, enzyme therapeutics can also target other molecules in the extracellular environment. In the clinical setting, the primary example is the use of amino acid–degrading enzymes as an anticancer strategy. Unlike healthy cells, rapidly growing tumor cells may be auxotrophic for certain metabolites, whose depletion in the plasma can selectively inhibit tumor growth. For example, lymphoid tumors are auxotrophic for asparagine because they lack asparagine synthetase activity, and recombinant, PEGylated L-asparaginase (Oncaspar®, Enzon) is an FDA-approved leukemia treatment (Pasut et al. 2008). A second enzyme, PEG-arginine deiminase (ADI-PEG 20, Pheonix Pharmacologics), is currently undergoing clinical trials for the treatment of the arginine-auxotrophic tumors melanoma (clinical trial phase I/II completed) and hepatocellular carcinoma (clinical trial phase II/III in progress) (Ni et al. 2008). Recent work indicates that renal cell carcinoma, a cancer that often metastasizes and is notoriously difficult to treat by conventional therapies, may also be susceptible to arginine deprivation [51]. A diverse group of solid tumor types, including lung, prostate, and bladder tumors are methionine-dependent (Mecham et al. 1983), and PEGylated recombinant methioninase has been studied as a cancer therapy in animal models (Pasut et al. 2008). Notably, the aforementioned enzymes are immunogenic, likely because they originate from bacterial sources, and as a result PEGylation was needed to reduce immunogenicity and prolong serum half-life (Pasut et al. 2008). Finally, PEGylated recombinant human arginase has been proposed as a treatment of arginine-dependent hepatocellular carcinoma resistant to PEG-arginine deiminase and is currently undergoing preclinical study [53].

Enzyme therapeutics have also found use in the treatment of cystic fibrosis. Frequent bacterial infections lead to accumulation and eventual lysis of neutrophils in the lungs, releasing extracellular DNA to form abnormally viscous mucus. Dornase alfa, or recombinant human DNase I (Pulmozyme®, Genentech), is delivered to the lungs as an aerosol and degrades extracellular DNA to improve lung function, quality of life, and prevent exacerbations of disease (Thomson 1995). Dornase alfa is produced by CHO cell culture, resulting in a glycosylated protein with minimized immunogenicity, but this also contributes to the high cost of the drug.
Protein Therapeutics as Replacements for Defective or Deficient Proteins

Protein Hormones

Some medical conditions characterized by deficiency or complete loss of an endogenous protein can be treated by replacement therapy. The most widely recognized protein therapeutics in this area are the protein hormones, including insulin and human growth hormone. Tens of millions of individuals worldwide suffer from Type I and Type II diabetes, and many are dependent on injection of insulin, a peptide consisting of one 21-amino-acid chain (A chain) and a separate 31-amino-acid chain (B chain) linked by two disulfide bonds. Extensive molecular engineering of this simple molecule has resulted in numerous short- and long-acting insulin analogs that are used to control mealtime glucose spikes and meet the body’s basal insulin need, respectively. These analogs can differ in terms of primary amino-acid sequence or may be chemically modified. For example, long-acting insulin glargine (Lantus®, Aventis) differs from human insulin by a substitution of asparagine for glycine at position 21 on the A chain and by the addition of two arginine residues to the C-terminus of the B chain (Sadrzadeh et al. 2007). Diarginyl insulin, possessing two extra arginines on the B chain, occurs naturally at low concentrations. A shifted isoelectric point makes this molecule soluble at mildly acidic pH, but it precipitates under physiological conditions; when injected, most diarginyl insulin degrades in subcutaneous tissue before it can be absorbed (Home and Ashwell 2002). The amino-acid substitution at position 21 enhances stability and increases bioavailability, providing a steady release of active insulin with only once-a-day injection (Bahr et al. 1997; Home and Ashwell 2002). Other long-acting forms of insulin were engineered by conjugation with fatty acids (Levemir®, Novo Nordisk) and PEGylation (InsuLAR, PR Pharmaceuticals) (Sadrzadeh et al. 2007). Alternatively, in fast-acting insulin lispro (Humalog®, Eli Lilly), the lysine at position 28 and the proline at position 29 of the B chain are inverted (Sadrzadeh et al. 2007). These mutations originated from the related protein hormone insulin-like growth factor I. Due to variation in the C-terminal ends of their respective B chains, natural insulin forms dimers and hexamers that slow absorption after subcutaneous injection, while insulin-like growth factor I does not self-associate (Holleman and Hoekstra 1997). By reversing residues 28 and 29, the self-association of insulin is abolished due to steric hindrance (Holleman and Hoekstra 1997). Injected insulin monomers absorb rapidly with a short duration of activity to control the spike in glucose associated with meals (Eckardt and Eckel 2008).

The second well-known protein hormone, recombinant human growth hormone (rhGH), has been available for several decades for the treatment of growth hormone deficiency in children and adults. rhGH also finds use in treating growth failure in children caused by other disorders (Bajpai and Menon 2005), as well as in treatment of HIV-associated wasting and lipodystrophy [61]. Current manufacturing processes generate two pharmacologically equivalent products: a 191-amino-acid protein identical to the natural human growth hormone or a 192–amino-acid protein possessing an additional N-terminal methionine [62]. The short plasma half-life of rhGH forces daily injection to maximize therapeutic benefit, making compliance an issue, particularly for pediatric patients. A sustained-release injection was developed via the collaboration of Genentech, Inc.
and Akermes, Inc., but commercialization of this product, Nutropin Depot®, was discontinued in 2004. Recently, conjugation of human growth hormone to human serum albumin [63] and polyethylene glycol (Cox et al. 2007) has been studied as a means to create more stable, longer-acting formulations for less frequent injections. Further, computational design of human growth hormone was carried out by Dahiyat and colleagues to create thermostabilized rhGH mutants (Filikov et al. 2002). To lower the potential for immunogenicity, analysis was focused on the core of the protein rather than its surface, and resulted in mutants with improved van der Waals interaction, hydrophobic substitutions at polar serine and threonine residues, and better burial of hydrophobic groups (Filikov et al. 2002). Higher thermostability may improve the stability of rhGH in the body, thereby increasing its half-life.

**Coagulation Factors**

Management of acquired or genetic loss of coagulation factors is another long-standing field of protein replacement therapy, and transfusion of human plasma–derived coagulation factors is the traditional treatment of bleeding disorders. Unfortunately, the use of human-derived products exposes transfusion recipients to risk of infection, and in the 1970s and 1980s, contaminated clotting factor concentrates led to HIV infection in nearly half of the United States’ population of hemophiliacs (Key and Negrier 2007). The sequencing of genes encoding human coagulation factors, as well as advances in molecular biology techniques and mammalian cell culture, has made recombinant production a reality (Pipe 2005). The availability of safe, pathogen-free coagulation factors has also created the opportunity to explore these therapies in “off-label” uses in nonhemophiliac patients. For example, while plasma-derived protein C is used to treat inherited protein C deficiency, recombinant activated protein C has been investigated in treatment of sepsis, an often fatal condition that leads to a rapid depletion of protein C and limited protein C activation (Key and Negrier 2007). Reasoning that the ability of activated protein C to reduce mortality in septic patients is related to its anti-inflammatory and anti-apoptotic activity but unrelated to its anticoagulant activity, researchers used site-directed mutagenesis to engineer a mutant protein C with reduced coagulant activity but normal anti-apoptotic activity (Mosnier et al. 2004). In the future, this mutant may prove effective at treating activated protein C depletion during sepsis without leading to severe bleeding complications (Mosnier et al. 2007).

**Enzyme Replacement Therapy**

The third category is enzyme replacement therapy (ERT) to treat the acquired or hereditary loss of an enzyme. ERT has revolutionized the treatment of several rare genetic diseases, including lysosomal storage disorders. In this case, successful therapy is contingent upon delivering exogenously supplied replacement enzyme to the intracellular lysosome and targeting specific cell types for maximum effectiveness. Gaucher disease ERT (Cerezyme®, Genzyme Corp) is accomplished by post-translational modification of the replacement enzyme β-Glucocerebrosidase (GCase). Glycosylated GCase is produced in CHO cells, and sugar residues are subsequently removed from the carbohydrate chains of the glycosylated protein to generate a
terminal mannose (Beck 2007). These mannose-terminated glycosylation residues interact with the mannose receptor, which is restricted to the macrophage plasma membrane, to deliver GCase specifically to macrophage lysosomes (Grabowski and Hopkin 2003).

**Cytokines and Their Receptors as Protein Therapeutics**

Cytokines, regulatory proteins secreted by white blood cells and a number of other cell types, and their receptors are important both as targets and tools in designing protein therapeutics. The term cytokine encompasses a variety of endogenous proteins, including colony stimulating factors (CSFs), epidermal growth factors (EGFs), interleukins, interferons, tumor necrosis factors, and others. It is well understood that cytokines can both protect from and contribute to disease, and accordingly, both recombinant cytokines and anticytokine antibodies or soluble receptors find clinical use (Vilcek and Feldmann 2004).

**Colony Stimulating Factors**

The most widely recognized cytokine therapeutic, owing to its misuse as a performance-enhancing drug among professional athletes, is likely erythropoietin, a CSF that increases proliferation of red blood cells (Segura et al. 2007). In the clinical setting, recombinant human erythropoietin (Epogen®, Amgen Inc; Procrit, Ortho Biotech Products, LP) treats anemia in patients undergoing chemotherapy or suffering from renal failure. Erythropoietin can also reduce the requirement for blood transfusion during surgery. AraNESP® (Amgen, Inc), or Darbepoetin alfa, was glycoengineered to contain five N-linked, sialic acid–containing carbohydrate chains, two more than recombinant human erythropoietin, which provided approximately threefold higher serum half-life, and therefore can be administered less often (Egrie and Browne 2001). Two other recombinant CSFs, granulocyte CSF (Neupogen and Neulasta, PEGylated, Amgen Inc) and granulocyte-macrophage CSF (Leucomax, Novartis, and others) are approved for use in patients undergoing chemotherapy or following a bone marrow transplant (Vilcek and Feldmann 2004). Recent in vivo pharmacokinetic studies of a mono-PEGylated recombinant human granulocyte CSF in rodents showed up to 40-fold enhanced serum half-life (Lee et al. 2008). Granulocyte CSF was also engineered via computational design to incorporate beneficial histidines on the binding surface of the protein. Sites were chosen such that under physiological conditions, a histidine of neutral charge does not adversely affect binding to the receptor, but in the slightly acid endosome, protonation results in a positive charge that disrupts receptor binding (Sarkar et al. 2002). In the endosome, weaker binding of granulocyte CSF to its receptor favors release of the cytokine and its receptor on the surface (recycling) over trafficking to the lysosome and subsequent degradation; as a result, improved potency and half-life were observed (Sarkar et al. 2002).

**Interleukins and Their Receptors**

Two recombinant human interleukins (ILs) are currently approved for therapeutic use. The first is a modified recombinant human IL-2 produced in *E. coli*. This
IL-2 analog, Aldesleukin (Proleukin®, Chiron Corp), is not glycosylated, lacks an N-terminal alanine, and possesses a cysteine-to-serine mutation at residue number 125. Proleukin® is FDA-approved for treatment of both metastatic renal cell carcinoma and metastatic melanoma and is employed as a standard treatment either alone or in combination with chemotherapy or interferon-α (Dillman 1999). However, the high systemic doses of IL-2 used in treatment of these cancers also leads to significant toxicity due to activation of natural killer (NK) cells (Atkins 2002). To reduce toxicity and enhance tumor killing, fusions of IL-2 with single chain T-cell receptor [79] or antibodies (Gillies et al. 1992; Lee et al. 2006) have been tested as a means to target IL-2 to tumor cells. Additionally, genetic engineering was used to enhance the affinity of IL-2 to a subunit of its receptor (IL-2Rα) found on human T cells but not NK cells (Rao et al. 2003). Higher affinity IL-2 mutants were isolated by yeast surface display and were shown to activate human T cells more potently than were wild type IL-2 (Rao et al. 2005). Improved activation of T cells may allow for lower dosage of recombinant IL-2 in the clinic, thereby reducing toxicity. As an added benefit, tight binding to the IL-2 receptors on activated human T cells essentially sequesters these IL-2 mutants so that they cannot interact with IL-2 receptors on NK cells (Rao et al. 2004).

The IL-2 receptor is also a useful clinical target. Fusion of IL-2 to toxins has been used to achieve targeted killing of IL-2 receptor-expressing cells. Specifically, a fusion of IL-2 and diphtheria toxin, Denileukin diftitox (Ontak®, Seragen Inc), is approved for treatment of persistent or recurrent cutaneous T-cell lymphoma in patients whose malignant cells express the CD25 component of the IL-2 receptor (Wong et al. 2007). Recently, a fusion of modified aerolysin and IL-2 was reported. Aerolysin is a bacterial toxin that binds to glycosylphosphatidylinositol-anchored proteins on the cell membrane and oligomerizes to create channels that trigger cell death. A variant of aerolysin with reduced binding to its native target was fused to IL-2 to create a chimera active only against cells expressing the IL-2 receptor [86]. This IL-2 fusion is a promising candidate for therapeutic use, and a fusion of modified aerolysin to prostate-specific antigen for the treatment of localized prostate cancer is already undergoing phase I clinical trials [86, 87]. Further, two monoclonal antibodies against the IL-2 receptor, basiliximab (Simulect®, Novartis Pharmaceuticals) and daclizumab (Zenapax®, Hoffmann-La Roche Inc) are used to prevent rejection following organ transplantation (Van Gelder et al. 2004).

The second interleukin approved for clinical application is Oprelvekin (Neumega®, Genetics Institute Inc), a nonglycosylated recombinant IL-11 produced in E. coli, which is also missing an N-terminal proline (Dorner et al. 1997). Oprelvekin prevents severe thrombocytopenia and reduces platelet transfusion requirements following myelosuppressive chemotherapy and bone marrow failure [90, 91]. Recently, PEGylation was reported to prolong serum half-life and enhance pharmacological efficacy of recombinant human IL-11, which may allow for less frequent injection [92].

**Interferons**

The interferons are a family of secreted proteins that regulate resistance to viral infections, enhance innate and acquired immune responses, and modulate normal and tumor cell survival (Borden et al. 2007). A number of interferon therapeutics
are approved for human use. Interferon-α (IFN-α) is approved both as the purified natural human protein and as the nonglycosylated recombinant proteins IFN-α-2a and IFN-α-2b, which differ by a substitution at amino acid position 23 (Roferon A®, Hoffmann-La Roche Inc; Intron A®, Schering Corp). The advantages of PEGylation led to the availability of both of these as PEG conjugates (Pegasys®, Hoffman-La Roche Inc; PEG-Intron®, Schering Corp). IFN-α and PEGylated IFN-α are standard treatments for chronic Hepatitis B and C and have been applied in the treatment of a number of different viral infections (Borden et al. 2007).

Variants of IFN-α have also been created using genetic engineering techniques. A novel IFN-α protein, termed consensus interferon (CIFN), was created by assigning the most commonly encountered amino acids at each position of several IFN-α nonallelic subtypes to create a consensus sequence (Keeffe and Hollinger 1997). Also known as Interferon alfacon-1 (Infergen®, InterMune Inc), CIFN was shown to be more potent than natural IFN-α \textit{in vitro} and is currently approved for clinical use (Alberti 1999). In 2007, DNA shuffling was used to create new IFN-α variants with higher antiviral activity than IFN-α-2a, IFN-α-2b, and CIFN, and importantly, decreased antiproliferative activity [96]. The antiproliferative activity of IFN-α is helpful in the treatment of cancers, but in the context of chronic viral illness, it may lead to bone marrow suppression and therefore limits the dosage used. The IFN-α mutants obtained in this study are thus “tailored” for treatment of viral infections due to their higher ration of antiviral to antiproliferative activity. The genetic engineering strategies applied to IFN-α are summarized in Figure 7.2. Additionally, to improve the serum half-life of IFN-α, glycoengineering was employed to create new, heavily glycosylated IFN-α analogs (Ceaglio et al. 2008). A fusion of IFN-α to human albumin was created for the same purpose, and this recombinant fusion protein is undergoing phase III clinical trials (Subramanian et al. 2007).

\textbf{FIGURE 7.2} Comparison of genetic engineering strategies used to create improved interferon-α mutants. The consensus method (left) was used to create highly active consensus interferon (CIFN), while DNA shuffling (right) was used to create interferon-α mutants tailored for treatment of chronic viral illness.
NEW CLASSES OF THERAPEUTIC PROTEINS UNDER DEVELOPMENT

T cells play a central role in cell-mediated immunity; however, the molecular mechanism underlying TCR recognition was not well understood until the late 1980s. Unlike antibodies that can recognize pathogens or their toxins directly, T cells recognize only short peptides derived from pathogens in complex with MHC molecules on the surface of antigen-presenting cells (APCs) through their TCRs. The nature of the interaction between TCRs and peptide-MHC (pMHC) complexes determines the function of the induced cellular immune responses. Therefore, both TCRs and MHC molecules can be potentially used in protein therapeutics (see the section titled “Targets of Protein Therapeutics and Modes of Action,” in this chapter).

MAJOR HISTOCOMPATIBILITY CLASS (MHC) PROTEINS

Malfunction of the adaptive immune system, consisting of class I (MHCI) and class II (MHCII) MHC proteins, has been implicated in many diseases, such as malaria, rheumatoid arthritis, type-I diabetes, and graft rejection. This has spurred great interest in developing MHC-based immunotherapeutics and immunodiagnostics methods. The development of pMHC tetramer, a multimeric form of peptide-MHC complexes, has revolutionized the field of T cell research. It enables direct detection and identification of antigen-specific T cells, modulation of T cell responses in vivo to treat graft rejection and autoimmune diseases, and detailed monitoring of cellular immune responses induced by immunotherapy, which is critical for a better understanding of tumor immunology and improved immune-based therapies. However, there are some limitations of pMHC tetramers that stem from their difficult recombinant production and the low affinity of the pMHC monomer. Therefore, it is highly desirable to engineer MHC molecules with improved solubility or higher TCR-binding affinity.

Phage Display

Phage display of MHCs is challenging, because MHC molecules are large, heterodimeric membrane proteins with high glycosylation and multiple disulfide bonds. In addition, in the absence of their transmembrane domains, their α and β polypeptide chains are unable to assemble properly and tend to aggregate. To date, there are only three successful examples of MHCI phage display, with the first one reported in 2000 (Le Doussal et al. 2000; Vest Hansen et al. 2001; Kurokawa et al. 2002). Although the displayed pMHC complexes were correctly folded and capable of binding specific antigenic peptides, no significant interaction with relevant T cells was detected. Therefore, it is necessary to further optimize and develop novel design of the pMHC phage display system for efficient and stable T cell recognition.

Cell Surface Display

Compared to phage display, yeast surface display is more effective for the display of MHC proteins and has been used to express both MHCI (Brophy et al. 2003; Jones et al. 2006) and MHCII (Starwalt et al. 2003; Esteban and Zhao 2004; Boder et al.
2005; Wen et al. 2008) proteins. Functional display of a mutant single-chain murine MHCI protein was evidenced not only by recognition of conformation-specific antibodies but also by direct binding of a specific TCR that has been engineered to have high affinity (see the section titled “T Cell Receptors (TCRs),” in this chapter) (Brophy et al. 2003). More significantly, yeast cells displaying pMHC complexes upregulated the surface expression level of an early activation marker on naive T cells isolated from mice. Although the authors did not rule out the possibility of T cell autostimulation, this study clearly suggested that yeast display could be used for directed evolution of pMHC complexes. Indeed, the same group later successfully isolated stabilized mutants of a single-chain murine MHCII protein (which is known to be unstable and difficult to work with) from either a focused library created by site-directed mutagenesis or a library created by random mutagenesis (Starwalt et al. 2003). In a similar study, mutants of a single-chain human MHCII protein without a covalently attached peptide were stably displayed on yeast cell surface (Esteban and Zhao 2004). These MHCII proteins exhibited specific and fast peptide-binding kinetics, as well as high thermostability. Interestingly, although the single-chain gene construct did not include any peptide, the peptide-binding groove of the displayed MHCII mutants was not empty, but occupied presumably by yeast endogenous peptides, indicating the importance of binding peptides in stabilizing MHCII expression. By incorporating an MHCII-binding peptide in the single-chain construct, the pMHC complex was functionally displayed on yeast cell surface without introduction of any mutations and capable of activating immobilized hybridoma T cells (Wen et al. 2008). More importantly, the authors demonstrated that yeast display could be used in combination with expression cloning to identify T cell epitopes from a pathogen-derived peptide library.

Recently, baculovirus-infected insect cells have also been used for MHC display (Crawford et al. 2004; Wang et al. 2005; Crawford et al. 2006). As higher-order eukaryotic cells, insect cells have fewer problems with MHC expression compared to yeast. So far, baculovirus display of MHC has been used only to identify T cell epitopes/mimotopes from peptide libraries. Nevertheless, the ability of insect cell–displayed pMHC complexes to directly activate relevant T cells (Crawford et al. 2004) makes this system a promising platform for pMHC engineering.

**T Cell Receptors (TCRs)**

TCRs are a pivotal element in almost every aspect of T lymphocytes, including their development, proliferation, differentiation, activity, and specificity. Therefore, researchers are exploring the potential of soluble TCRs to be used as immunotherapeutic or immunodiagnostic reagents to specifically target the pMHC complex (Miles et al. 2006), just as antibodies are used to neutralize or opsonize their antigens. However, there are several obstacles impeding therapeutic applications of TCRs, including difficult recombinant production, instability, and low pMHC binding affinity. Therefore, efforts have been devoted to the engineering and design of soluble, stable, and high-affinity TCRs.
Protein Fusions and Rational Design

Until recently, there was no generally applicable method of producing soluble TCRs (Molloy et al. 2005). Many of the initial strategies, such as removing exposed hydrophobic residues, or fusion to antibody constant regions or thioredoxin (Andrews et al. 1996; Shusta et al. 1999), worked for a very limited number of TCRs. Later, Jun/Fos leucine zipper domains were introduced as fusions to the C-termini of the α/β TCR extracellular domains, respectively (Willcox et al. 1999). The incorporation of leucine zipper domains significantly stabilized the TCR, while maintaining its ligand specificity; however, it raised the potential of immunogenicity. Another generally applicable method involved introducing a non-native interchain disulfide bond, predicted by molecular modeling based on the TCR crystal structure, in the TCR invariant region (Boulter et al. 2003). The resulting disulfide-stabilized TCR (dsTCR) was highly stable, and the sequence/structural change was minimal compared to wild-type TCR, reducing its possibility of being immunogenic. More importantly, the dsTCR construct enabled phage display (Figure 7.3A), which represented a powerful directed evolution platform for TCR engineering (Li et al. 2005). By using the dsTCR format, ten different human class I– and class II–restricted TCRs were successfully displayed on phage particles (Li et al. 2005).

**Figure 7.3** TCR engineering using (A) phage display and (B) yeast display, adapted from (Boder and Wittrup 1997). A phage display library of TCRs is usually screened by panning phage particles on a surface/matrix coated with purified pMHC complexes, while a yeast display library can be rapidly screened by FACS using fluorophore-labeled (such as phycoerythrin, PE) pMHC complexes.
**Phage Display**

The highest affinity TCRs yet reported (with picomolar affinities for their specific pMHC ligand) were achieved using phage display and directed evolution (Li et al. 2005) (Figure 7.3A). By using random mutagenesis, human TCR mutants with pMHC affinities up to 26 pM were isolated, representing more than a million-fold improvement over the wild-type TCR. In addition, the half-life of pMHC binding was also improved by ~8000-fold to ~1000 minutes at 25°C. More significantly, these high affinity TCRs showed a high degree of antigen specificity and no cross-reactivity with endogenous pMHC complexes, and they enabled direct visualization of specific pMHC complexes on tumor cells for the first time (Purbhoo et al. 2006). When transfected into human T cells, the mutant TCR with highest pMHC affinity completely lost its antigen specificity, while those expressing TCRs with lower pMHC affinity (with $K_D$ values of 450 nM and 4 μM) responded in an antigen-specific manner (Zhao et al. 2007). This indicates that genetically engineered T cells with midrange pMHC affinity might be useful in immunotherapeutics.

**Yeast Surface Display**

As with MHC proteins, directed evolution and yeast surface display have also been applied to engineer soluble TCRs (Figure 7.3B). By using an *E. coli* mutator strain XL1-Red, FACS screening, and combining mutations, TCR mutants with increased thermal stability and secretion efficiency were identified (Kieke et al. 1999; Shusta et al. 1999). Selected mutations were combined and resulted in a TCR mutant that was stable for >1 hour at 65°C, had a solubility of over 4 mg/mL, and had a shake-flask expression level of 7.5 mg/L (Shusta et al. 2000). More importantly, although mutations were introduced, the resulting TCRs retained their ligand-binding specificity, making yeast display an attractive engineering platform for engineering TCRs with high affinity to pMHC (Kieke et al. 1999; Shusta et al. 2000). The first example of *in vitro* affinity maturation of a TCR was reported in 2000 using yeast surface display (Holler et al. 2000). A focused library of MHCI-restricted single-chain TCR was constructed by mutating the CDR3 (complementarity-determining region three) of the $\alpha$-chain, and mutants with greater than 100-fold higher pMHC binding affinity were identified. Unlike the wild-type TCR, the soluble monomeric form of the high-affinity TCR was capable of directly detecting specific pMHC complexes on APCs. *In vivo* studies showed that a mouse T cell hybridoma transfected with the high affinity TCR responded to a significantly lower concentration of antigenic peptide (Holler et al. 2001). In several follow-up studies, it was shown that mutations in the CDR1 and CDR2 regions could also contribute to improving the pMHC binding affinity of TCRs (Chlewicki et al. 2005; Weber et al. 2005).

**TCR-Like Antibodies**

Researchers have used phage display to create TCR-like antibodies, which recognize specific peptides in complex with MHC molecules (Andersen et al. 1996; Denkberg and Reiter 2006). There have been ~100 different TCR-like antibodies generated to date, recognizing antigens involved in several infectious diseases and cancer.
However, the in vivo targeting capability of these TCR-like antibodies has not yet been demonstrated. With further development, this new class of antibodies might lead to novel therapeutic and diagnostic solutions.

CONCLUSION AND FUTURE PROSPECTS

Protein therapeutics are a very important part of modern medicine, and in certain situations they are the only effective therapies. Expanding with an ever-increasing speed, the protein therapeutics market is projected to reach $87 billion by year 2010 (KaloramaInformation 2006). To realize such great potential, continuous efforts are required to optimize their efficacy, while simultaneously discovering novel protein drugs. As exemplified in this chapter, protein engineering and design have long been valuable tools in developing effective protein therapeutics with improved affinity, specificity, activity, stability, pharmacokinetics, pharmacodynamics, and reduced immunogenicity and improved productivity. In the future, with better understanding of protein structure-function relationships and rapid development of in silico bioinformatics and systems biology techniques, we are more likely to see an increased synergy between protein engineering and design strategies. In terms of discovering novel protein therapeutics, it is necessary to incorporate strengths from multiple disciplines, such as molecular biology, pathology, immunology, and nanotechnology. With orchestrated efforts, the “Golden Age” of protein therapeutics can be realized.

REFERENCES


