4.1 INTRODUCTION

The modern era of biological therapeutics may be identified with the FDA approval of recombinant insulin (Humulin) in 1982, produced in *E. coli*, and recombinant erythropoietin (EPO) in 1989 (Epogen); since glycosylation of EPO is essential to its function, it was necessarily produced in a mammalian cell line [a Chinese hamster ovary (CHO) cell line]. Despite extensive clinical experience, adverse reactions to these recombinant molecules are still encountered: ~2% for insulin (Ghazavi and Johnston, 2011) and rarer, but more devastating, for EPO (Macdougall et al., 2012). These incidences are frequently due to the patient developing antibodies that are specific for the therapeutic and neutralize its activity [antitherapeutic antibodies (ATA): antidrug antibodies (ADA)]. The development of ADA suggests that the therapeutic is being recognized as "foreign" (nonself) by the patient's immune system, owing to the presence of molecules that exhibit structural features different from those of the endogenous protein/glycoprotein (P/GP). Ironically, a high incidence of ADA is encountered for recombinant antibody therapeutics (mAbs), a sort of "good cop/bad cop" situation; the explanation lies in the fact that, by definition, each mAb exhibits a unique specificity due to its unique structure.

Currently, some 90 protein therapeutics have been established in the clinic, and patents for those approved early in the development of biologic drugs have expired or will do so imminently. The high costs of these drugs impose financial pressures on national and private health care bodies, with consequent encouragement for the development and prescribing of copies of innovator biotherapeutics (Biosimilars, EU; Follow-on biologics, FDA). While the availability of biosimilars reduces the "cost of goods" (CoG), advances in formulation and delivery will contribute to further reductions in "cost of treatment" (CoT). Additionally, genetic and glycosylation engineering is being embraced to generate innovator "biobetter" molecules that express "enhanced" attributes, compared to the currently approved biologics. Given the complexity of the human proteome and cellome, there would seem to be virtually unlimited scope for the continued development of recombinant P/GP therapeutics.

The starting point for the generation of recombinant P/GPs is extensive characterization of the structure and function of the endogenous (natural) molecule. This is not a trivial exercise, for while the gene sequence determines the primary amino acid sequence, it does not provide a guide to the precise structure of the active molecule. Additional parameters include the *in vivo*: conformation (secondary, tertiary, and quaternary), chemical properties (charge, hydrophilicity, and hydrophobicity), post-translational (PTM) and chemical modifications (CMs), and the microenvironment in which the molecule is functionally active. Consequently, the endogenous P/GP exhibits structural heterogeneity, and individual isoforms may influence the rate of protein turnover, determining whether or when a given molecule enters a degradative pathway (Welle, 1999). These initial considerations are compounded when one is attempting to establish the structure of a "native" P/GP *ex vivo* because additional heterogeneities may be introduced during its isolation, purification, and characterization.

This results in a "Catch 22" or biological "Heisenberg uncertainty principle" situation; the purer the isolated protein, the lower the yield and the less certain one can