

Post-translational Modifications of Recombinant Proteins: Significance for Biopharmaceuticals

Nigel Jenkins · Lisa Murphy · Ray Tyther

Published online: 8 March 2008
© Humana Press 2008

Abstract The production of recombinant therapeutic proteins is one of the fastest growing sectors of the pharmaceutical industry, particularly monoclonal antibodies and Fc-fusion proteins. Currently, mammalian cells are the dominant production system for these proteins because they can perform complex post-translational modifications that are often required for efficient secretion, drug efficacy, and stability. These protein modifications include misfolding and aggregation, oxidation of methionine, deamidation of asparagine and glutamine, variable glycosylation, and proteolysis. Such modifications not only pose challenges for accurate and consistent bioprocessing, but also may have consequences for the patient in that incorrect modifications and aggregation can lead to an immune response to the therapeutic protein. This mini-review describes examples analytical and preventative advances in the fields of protein oxidation, deamidation, misfolding and aggregation (glycosylation is covered in other articles in this issue). The feasibility of partially replacing traditional analytical methods such as peptide mapping with high-throughput screens and their use in clone and media selection are evaluated. This review also discusses how further technical advances could improve the manufacturability, potency, and safety of biotherapeutics.

Keywords Recombinant protein · Cell culture · Aggregation · Folding · Oxidation · Deamidation · Screening · Therapeutics

Introduction

The ability to perform complex post-translational modifications is a major attribute that distinguishes mammalian cells from other production platforms used to produce biopharmaceuticals. Indeed, only a few biopharmaceutical proteins such as albumin (Recombunin) and insulin (Lispro) undergo simple modifications such that they can be manufactured using yeast or bacteria [1]. The most prevalent modifications include variable glycosylation, misfolding and aggregation, oxidation of methionine, deamidation of asparagine and glutamine, and proteolysis. Detecting and preventing these modifications has become a major challenge for the biotechnology industry. As the market for therapeutic protein production from mammalian cells has expanded, so too has the requirement for improved efficiency and stability of production [2]. Expression systems that exploit multiple gene copies or use strong promoters to increase intracellular mRNA levels do not consistently increase the amount of secreted recombinant protein, in part due to bottlenecks in the secretory pathways [3].

Protein Misfolding and Aggregation

The folding and initial glycosylation of most secreted proteins take place in the endoplasmic reticulum (ER) lumen. However, misfolded proteins can accumulate as intracellular aggregates and induce dilation of the ER [4]. Therefore, the rate-limiting step for recombinant protein secretion becomes the exit of correctly folded polypeptide chain from the ER. Molecular chaperones such as heavy chain-binding protein (BiP) facilitate protein folding at high concentrations by binding of unfolded protein chains, prevention of aggregation, and/or support of refolding.

N. Jenkins (✉) · L. Murphy · R. Tyther
National Institute for Bioprocessing Research and Training (NIBRT), Engineering Building, University College Dublin, Belfield, Dublin 4, Ireland
e-mail: Nigel.Jenkins@NIBRT.ie

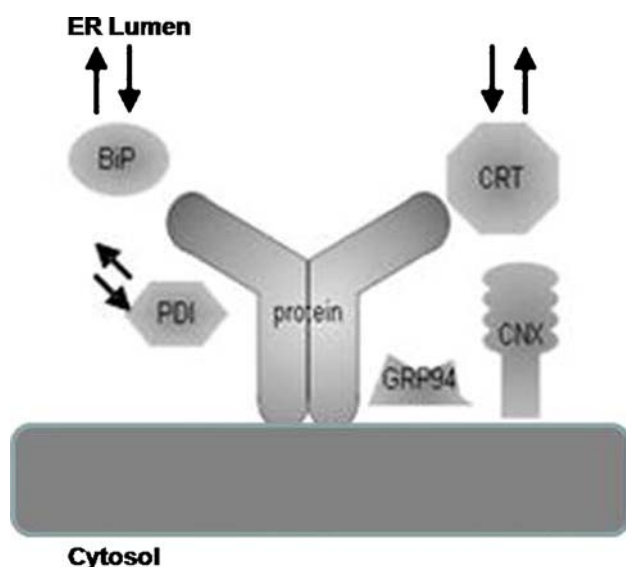


Fig. 1 Protein folding in the endoplasmic reticulum (ER). Secreted proteins are co-translationally located in the ER lumen. Both co-translational and post-translational folding events are chaperoned by ER-resident folding chaperones such as heavy chain binding protein (BiP), protein disulfide isomerase (PDI), glucose regulated protein 94 (GRP94), calnexin (CNX), and calreticulin (CRT). Disulfide bond formation is catalyzed by PDI and its accessory proteins. The chaperones and folding enzymes release the proteins from the ER only when they have folded correctly and assembled into oligomers

These and other molecules deliver misfolded and unfolded proteins to the unfolded protein pathway that degrades peptides in the endosome [5], as shown in Fig. 1. Additionally, chaperones such as calreticulin and calnexin provide another quality control mechanism by transient binding to newly synthesized glycoprotein intermediates, ensuring that only correctly folded proteins are released from the ER [6]. However, in high-producing cell lines the capacity of these chaperones can become overloaded with recombinant protein, leading to accumulation of misfolded and aggregated proteins, for example CHO cells making Antithrombin III (AT-III) [7]. In this study, some of the intracellular AT-III was found to form disulfide-bonded high molecular weight aggregates at expression levels that were rate-limiting to the secretion of this molecule ($7.4 \mu\text{g}/10^6$ cells/day). The data indicated that a portion of the AT-III was aggregated either as dimers or higher molecular weight forms, and that most of the aggregates contain disulfide bonds because they were not detected under reducing conditions. More AT-III aggregation was also observed in cells that had reached the late exponential phase of culture compared to the earlier exponential growth phase of culture, however, a metabolic trigger for this increase in aggregation could not be found.

Protein aggregates can arise by several mechanisms that include reversible or irreversible reactions, non-covalent

interactions between hydrophobic domains, or the formation of disulfide bonds. Some are insoluble and others remain in solution. For non-vaccine biotherapeutics, all types of aggregates are considered undesirable since small soluble aggregates may become immunogenic and larger particulates may cause problems at the site of administration [8]. Whilst there are US Pharmacopoeia guidelines for the number of particles of size ≥ 10 and $\geq 25 \mu\text{m}$ that are acceptable in a pharmaceutical preparation, the permissible number of soluble aggregates are not well defined.

Disulfide bonds result from the coupling of unpaired thiols on Cysteine residues, whilst this is essential for the correct formation of some multimeric proteins such as IgG it can also be the source of covalent aggregate formation and protein misfolding. One effective way of reducing the free thiol content and increasing disulfide bond formation of recombinant IgG produced in CHO cells is to add low amounts (up to $100 \mu\text{M}$) of the oxidizing agent copper sulfate. This resulted in a 10-fold reduction of percentage free thiols, and significant (3-fold) reductions were observed with as little as $5 \mu\text{M}$ copper sulfate [9]. In order to assist the formation of disulfide bonds between thiol groups, the lumen of the ER provides an oxidizing environment, the substrate glutathione and enzymes such as protein disulfide isomerase (PDI), which catalyze the formation of protein tertiary structures [10]. PDI catalyzes the formation and breakage of disulfide bonds between cysteine residues within proteins as they fold, allowing proteins to quickly find the correct arrangement of disulfide bonds.

Effects of Protein Aggregation

Evidence has been found that protein aggregates can continue to form in the supernatant after cells (producing recombinant IgG) have been harvested [8], and the aggregation problem is not confined to upstream bioprocessing. For example, the techniques used to inactivate viruses during downstream processing such as exposure to detergents or extremes of pH can inadvertently damage and aggregate the protein product [11]. Low pH conditions (pH 2–4) are also commonly used to elute antibodies from Protein A capture columns during purification. Multiple filtration steps are also used in protein purification for concentration, buffer exchange, and virus removal. Large protein aggregates can cause membrane fouling and the high pressures employed may also increase aggregation during these process steps. Excipients such as sugars [12] and arginine [13] are often used to suppress aggregate formation during protein purification and formulation.

In addition to aggregates causing manufacturing problems, the administration of proteins with low-level aggregate contamination can lead to an immune response in the patient,

resulting in inhibitory antibodies to the therapeutic protein [14]. This can be a particular problem with multimeric proteins such as recombinant IgG and blood clotting proteins such as Factor VIII [15]. Another example is where the biotherapeutic forms aggregates with its excipient: a formulation of interferon alpha-2a became oxidized at room temperature and formed aggregates with the excipient human serum albumin. This induced an immune response to IFN-alpha-2a, and changing to a liquid, albumin-free formulation stored at 4°C reduced the immune reaction [16].

Cell Line Engineering to Improve Protein Quality

Engineering of chaperone systems by overexpressing a single component of the ER-resident protein folding machinery has overall yielded mixed results. BiP overexpression has been shown to increase heterologous protein secretion for about half of the heterologous proteins studied [17] and in some cases inhibits protein secretion [18]. It has been proposed that either insufficient ATP levels or the lack of co-chaperones such as Lhs1p may become rate-limiting to BiP functions. In addition, increased BiP activity may stall the GRP94, calnexin–calreticulin chaperone and unfolded protein response machineries, because of the hierarchy of ER luminal chaperone systems.

As with BiP, PDI overexpression does not always lead to improved levels of protein secretion [19, 20]. Regeneration of oxidized PDI is essential for consistent disulfide bond formation, and other factors in this pathway such as glutathione availability and the activity of Ero1p (the

enzyme that oxidizes PDI) may become rate-limiting in PDI over-expressing cells [21].

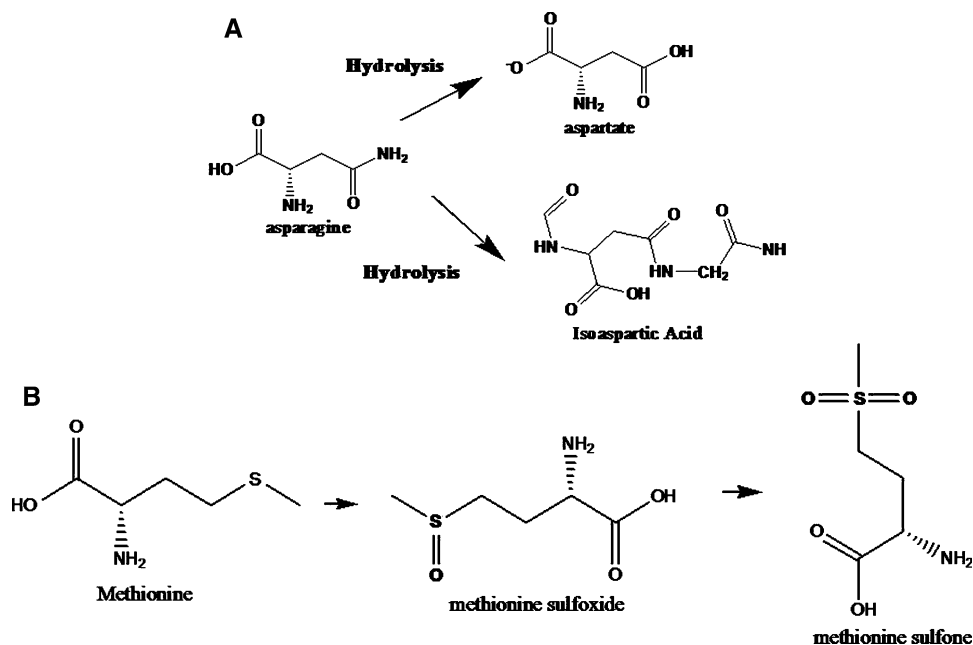
Clearly, more work is required to optimize the secretory and folding capacities of recombinant mammalian cells both to improve yields and increase product quality. Recent advances in CHO genomics [22] should facilitate these approaches.

Oxidation and Deamidation

Some of the methionine residues in a protein can be oxidized to methionine sulfoxide or even methionine sulfone. The exposure of susceptible residues on the cell surface can determine the extent of methionine oxidation, and conditions such as the cell culture media components and formulation excipients can be used to protect the protein from oxidative damage [23]. In cases such as α_1 -antitrypsin (used in the treatment of emphysema) oxidation of methionine residues leads to a loss of its critical anti-elastase activity [24].

Deamidation of asparagine residue to form aspartic acid and iso-aspartic acid is another cause of protein degradation, particularly during long-term storage (Fig. 2) [25]. Most of the deamidated asparagine forms iso-aspartate, which is not a natural amino acid and can potentially be immunogenic. This isomerization is accompanied by partial racemization of newly created isoaspartyl and aspartyl residues. This degradative event can have serious implications for drug efficacy, because the potency of one commercially available antibody was reduced to 70%

Fig. 2 Two types of degradative post-translational modifications: (a) asparagine deamidation and (b) methionine oxidation



post-deamidation [26], and deamidation has also been shown to alter the activity of stem cell factors [27]. The non-enzymic deamidation reaction is accelerated at alkaline pH, with surface asparagine residues being more susceptible. Glutamine residues can also be deamidated, but this reaction is one hundred times slower than asparagine deamidation and is rarely detected in recombinant proteins. Predicting which asparagine residues will be prone to deamidation in a new therapeutic protein can also be difficult.

Analytical Methods

Detecting protein aggregates can be a challenge, not least because aggregates vary in size from simple dimers to large particles several μm in diameter. Size-exclusion HPLC (SEC-HPLC) has become one of the standard techniques used to detect aggregates, but some artifacts may be introduced under the high pressures needed to run HPLC, and the large aggregates may not enter the column. A certain degree of incomplete product recovery is also evident with SEC due to protein adsorption to the gel matrix. Because of these limitations, the FDA usually requests orthogonal analytical techniques to confirm SEC-HPLC results. Analytical ultracentrifugation (AUC) has been proposed as an alternative detection system based on differences in sedimentation rates, but the AUC equipment can be quite expensive [28]. Because AUC is both time-consuming and expensive, it may not lend itself to routine product characterization but rather serves as a validation/calibration technique. Larger aggregate particles can be detected using laser light scattering [29].

Protein aggregate-sensitive dyes may help by complementing existing detection methodologies used in bioprocessing. Among the most widely used dyes is Congo Red, a sulfonated azo dye that undergoes a spectral shift from 490 to 530 nm when bound to fibrils and other forms of protein aggregates [30]. In addition to the parent molecule, various Congo Red derivatives are available which may differ in their lipophilicity or ability to undergo fluorescence [31]. The other principal dye utilized in detecting aggregates is thioflavin T, a benzothiolate dye that exhibits enhanced fluorescence at 480 nm when bound to amyloid. Thioflavin T has also been adapted for fluorescent spectrophotometric assay [32].

Various other dyes such as 4-(dicyanovinyl)-julolidine (DCVJ) [31] have also been used in both static and kinetic assays to characterize aggregation events using both steady state and time-resolved fluorescence techniques. Introducing 96-well plate spectrophotometric assays based on these dyes could facilitate protein quality measurements during clone and media selection.

Total free thiol content can be rapidly assessed spectrophotometrically using Ellman's reagent or a fluorescent thiol-sensitive maleimide dye, such as Fluorescein-5-Maleimide, or by hydrophobic interaction chromatography [9]. Intra- or Inter-molecular disulfide bond formation can be detected using diagonal electrophoresis, whereby protein complexes are separated firstly undergo non-reducing then reducing conditions [33].

Methionine oxidation can be detected by mass spectrometry or UV detection of peptide fragments [34]. Peptides containing oxidized methionines can also be separated from natural fragments on a weak ion-exchange column [35] or by hydrophobic interaction chromatography [36].

Asparagine residues prone to deamidation can be identified through peptide mapping of the recombinant protein. This involves reduction, alkylation, and tryptic digestion of the protein and subsequent HPLC separation of the tryptic peptides prior to mass spectrometric analysis, and these multiple steps are difficult to adapt into high-throughput screens. Peptide mass fingerprinting has been employed to identify deamidation sites in human IgG [25]. The ISO-QUANT assay has been developed by Promega for detecting deamidation. This method exploits the ability of the enzyme L-isoaspartyl methyltransferase (PIMT) to transfer the active methyl group of S-adenosyl-L-methionine (SAM) onto the free alpha-carboxyl of iso-aspartate to form an O-methyl ester. The isoaspartyl methyl ester rapidly breaks down at neutral pH to form the same cyclic imide arising from deamidation or aspartate isomerization, with concomitant release of methanol. Incubating proteins with PIMT and a radioactive substrate (^3H -SAM), allows isoaspartyl sites to become quantitatively methylated and quantified by scintillation counting of ^3H -methanol. However, no detection method for oxidized methionines in proteins currently exists in a microtitre format that could be used for high-throughput screening.

Conclusions

The number of recombinant proteins in clinical trials for new and existing therapeutic targets continues to increase each year, as does the total amount of product required globally. It is also predicted that generic protein therapeutics will become more prevalent as the patents associated with the original proteins expire [37]. Despite the success of biotherapeutics there remain significant challenges to be overcome in maintaining product stability and efficacy throughout the production cycle and during long-term storage. These include the pathways reviewed in this article, as well as variable glycosylation, proteolytic cleavage and RNA splice variants. Advances in protein

analysis and a deeper knowledge of industrial process conditions should lead to preventative strategies that help counteract these degradative pathways and ensure product efficacy, safety and affordability.

Acknowledgments The authors wish to thank Prof. Martin Clynes and other members of the National Institute for Cellular Biotechnology (NICB), Dublin City University for the use of their laboratory facilities. The financial support of the Irish Industrial Development Agency (IDA, Ireland) is gratefully acknowledged.

References

- Roach, P., & Woodworth, J. R. (2002). Clinical pharmacokinetics and pharmacodynamics of insulin lispro mixtures. *Clinical Pharmacokinetics*, *41*, 1043–1057.
- Barnes, L. M., & Dickson, A. J. (2006). Mammalian cell factories for efficient and stable protein expression. *Current Opinion in Biotechnology*, *17*, 381–386.
- Schroder, M. K. C., & Friedl P. (1999). Quantitative analysis of transcription and translation in gene amplified Chinese hamster ovary cells on the basis of a kinetic model. *Cytotechnology*, *29*, 93–102.
- Schroder, M., & Kaufman, R. J. (2005). The mammalian unfolded protein response. *Annual Review of Biochemistry*, *74*, 739–789.
- Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., & Ron, D. (2000). Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biology*, *2*, 326–332.
- Molinari, M., & Helenius, A. (2000). Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science*, *288*, 331–333.
- Schroder, M., Schafer, R., & Friedl, P. (2002). Induction of protein aggregation in an early secretory compartment by elevation of expression level. *Biotechnology and Bioengineering*, *78*, 131–140.
- Cromwell, M. E., Hilario, E., & Jacobson, F. (2006). Protein aggregation and bioprocessing. *AAPS Journal*, *8*, E572–579.
- Chaderjian, W. B., Chin, E. T., Harris, R. J., & Etcheverry, T. M. (2005). Effect of copper sulfate on performance of a serum-free CHO cell culture process and the level of free thiol in the recombinant antibody expressed. *Biotechnology Progress*, *21*, 550–553.
- Chakravarthi, S., Jessop, C. E., & Bulleid, N. J. (2006). The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Report*, *7*, 271–275.
- Lin, J. J., Meyer, J. D., Carpenter, J. F., & Manning, M. C. (2000). Stability of human serum albumin during bioprocessing: Denaturation and aggregation during processing of albumin paste. *Pharmaceutical Research*, *17*, 391–396.
- Cleland, J. L., Lam, X., Kendrick, B., Yang, J., Yang, T. H., Overcashier, D., Brooks, D., Hsu, C., & Carpenter, J. F. (2001). A specific molar ratio of stabilizer to protein is required for storage stability of a lyophilized monoclonal antibody. *Journal of Pharmaceutical Sciences*, *90*, 310–321.
- Tsumoto, K., Ejima, D., Kita, Y., & Arakawa, T. (2005). Review: Why is arginine effective in suppressing aggregation? *Protein and Peptide Letters*, *12*, 613–619.
- Hermeling, S., Crommelin, D. J., Schellekens, H., & Jiskoot, W. (2004). Structure-immunogenicity relationships of therapeutic proteins. *Pharmaceutical Research*, *21*, 897–903.
- Purohit, V. S., Middaugh, C. R., & Balasubramanian, S. V. (2006). Influence of aggregation on immunogenicity of recombinant human Factor VIII in hemophilia A mice. *Journal of Pharmaceutical Sciences*, *95*, 358–371.
- Ryff, J. C. (1997). Clinical investigation of the immunogenicity of interferon-alpha 2a. *Journal of Interferon and Cytokine Research*, *17*(Suppl 1), S29–S33.
- Kleizen, B., & Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. *Current Opinion in Cell Biology*, *16*, 343–349.
- Dorner, A. J., Wasley, L. C., & Kaufman, R. J. (1992). Over-expression of GRP78 mitigates stress induction of glucose regulated proteins and blocks secretion of selective proteins in Chinese hamster ovary cells. *EMBO Journal*, *11*, 1563–1571.
- Butz, J., Niebauer, R. T., & Robinson, A. S. (2003). Co-expression of molecular chaperones does not improve the heterologous expression of mammalian G-protein coupled receptor expression in yeast. *Biotechnology and Bioengineering*, *84*, 292–304.
- Damasceno, L., Anderson, K. A., Ritter, G., Cregg, J. M., Old, L. J., & Batt, C. A. (2006). Co-overexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. *Applied Microbiology and Biotechnology*, *56*, 157–164.
- Chakravarthi, S., & Bulleid, N. J. (2004). Glutathione is required to regulate the formation of native disulfide bonds within proteins entering the secretory pathway. *Journal of Biological Chemistry*, *279*, 39872–39879.
- Wlaschin, K. F., Nissom, P. M., Gatti Mde, L., Ong, P. F., Arleen, S., Tan, K. S., Rink, A., Cham, B., Wong, K., Yap, M., et al. (2005). EST sequencing for gene discovery in Chinese hamster ovary cells. *Biotechnology and Bioengineering*, *91*, 592–606.
- Soenderkaer, S., Carpenter, J. F., van de Weert, M., Hansen, L. L., Flink, J., & Frokjaer, S. (2004). Effects of sucrose on rFVIIa aggregation and methionine oxidation. *European Journal of Pharmaceutical Sciences*, *21*, 597–606.
- Taggart, C., Cervantes-Laurean, D., Kim, G., McElvaney, N. G., Wehr, N., Moss, J., & Levine, R. L. (2000). Oxidation of either methionine 351 or methionine 358 in alpha 1-antitrypsin causes loss of anti-neutrophil elastase activity. *Journal of Biological Chemistry*, *275*, 27258–27265.
- Chelius, D., Rehder, D. S., & Bondarenko, P. V. (2005). Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. *Analytical Chemistry*, *77*, 6004–6011.
- Harris, R. J., Kabakoff, B., Macchi, F. D., Shen, F. J., Kwong, M., Andya, J. D., Shire, S. J., Bjork, N., Totpal, K., & Chen, A. B. (2001). Identification of multiple sources of charge heterogeneity in a recombinant antibody. *Journal of Chromatography B: Biomedical Science Applications*, *752*, 233–245.
- Hsu, Y. R., Chang, W. C., Mendiaz, E. A., Hara, S., Chow, D. T., Mann, M. B., Langley, K. E., & Lu, H. S. (1998). Selective deamidation of recombinant human stem cell factor during in vitro aging: Isolation and characterization of the aspartyl and isoaspartyl homodimers and heterodimers. *Biochemistry*, *37*, 2251–2262.
- Pekar, A., & Sukumar, M. (2007). Quantitation of aggregates in therapeutic proteins using sedimentation velocity analytical ultracentrifugation: Practical considerations that affect precision and accuracy. *Analytical Biochemistry*, *367*, 225–237.
- Ye, H. (2006). Simultaneous determination of protein aggregation, degradation, and absolute molecular weight by size exclusion chromatography-multiangle laser light scattering. *Analytical Biochemistry*, *356*, 76–85.
- Krol, M., Roterma, I., Piekarska, B., Konieczny, L., Rybarska, J., Stopa, B., Spolnik, P., & Szneler, E. (2005). An approach to understand the complexation of supramolecular dye Congo red with immunoglobulin L chain lambda. *Biopolymers*, *77*, 155–162.

31. Lindgren, M., Sorgjerd, K., & Hammarstrom, P. (2005). Detection and characterization of aggregates, prefibrillar amyloidogenic oligomers, and protofibrils using fluorescence spectroscopy. *Biophysical Journal*, *88*, 4200–4212.
32. Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Krishna, V., Grover, R. K., Roy, R., & Singh, S. (2005). Mechanism of thioflavin T binding to amyloid fibrils. *Journal of Structural Biology*, *151*, 229–238.
33. Lin, H. J., Lin, C. H., Tseng, H. C., & Chen, Y. H. (2006). Detecting disulfide crosslinks of high-molecular weight complexes in mouse SVS proteins by diagonal electrophoresis. *Analytical Biochemistry*, *352*, 296–298.
34. Houde, D., Kauppinen, P., Mhatre, R., & Lyubarskaya, Y. (2006). Determination of protein oxidation by mass spectrometry and method transfer to quality control. *Journal of Chromatography A*, *1123*, 189–198.
35. Chumsae, C., Gaza-Bulsecu, G., Sun, J., & Liu, H. (2006). Comparison of methionine oxidation in thermal stability and chemically stressed samples of a fully human monoclonal antibody. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *850*, 285–294.
36. Lam, X. M., Yang, J. Y., & Cleland, J. L. (1997). Antioxidants for prevention of methionine oxidation in recombinant monoclonal antibody HER2. *Journal of Pharmaceutical Sciences*, *86*, 1250–1255.
37. Schellekens, H., & Ryff, J. C. (2002). 'Biogenerics': The off-patent biotech products. *Trends in Pharmacological Sciences*, *23*, 119–121.