Characterization of Fc-Fusion Protein Aggregates Derived from Extracellular Domain Disulfide Bond Rearrangements

JAMES STRAND,[1] CHI-TING HUANG,[1] JIN XU[2]

[1]Acceleron Pharma, Cambridge, Massachusetts 02139

[2]Department of Chemistry, University of Massachusetts-Lowell, Lowell, Massachusetts 01854

Received 21 May 2012; revised 5 November 2012; accepted 28 November 2012

Published online 14 December 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.23421

ABSTRACT: Aggregation of protein biotherapeutics has consequences for decreasing production and has been implicated in immunogenicity. The mechanisms of protein aggregation vary depending on the protein and the expression system utilized, making it difficult to elucidate the conditions that promote their formation. Nonnative aggregation of recombinant immunoglobulin G protein therapeutics from mammalian expression systems has been extensively studied. To better understand the mechanisms behind aggregation of glycosylated fusion proteins produced in Chinese hamster ovarian cells, we have examined the high-molecular-weight (HMW) species of activin receptor-like kinase 1 Fc fusion protein. Size-exclusion chromatography and sodium dodecyl sulfate-polyacrylamide gel electrophoresis indicate that two populations of aggregate exist: (1) nondisulfide-linked, higher-order aggregates and (2) disulfide-linked oligomers. The largest aggregated species have increased nonnative structure, whereas the smallest aggregated species maintain structure similar to monomer. The HMW species display decreased levels of *O*-linked glycosylation, higher occupancy of high-mannose *N*-linked oligosaccharide structures, and overall less sialylation as their size increases. Disulfide-linked aggregate species were found to associate through the extracellular domain. *N*-linked glycosylation on the extracellular domain (ECD) appears to discourage disulfide-linked aggregation. Elucidation of the specific mechanisms behind disulfide-linked aggregate formation may assist in designing processes that limit aggregate formation in cell culture, with implications for increased production. © 2012 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci 102:441–453, 2013

Keywords: protein aggregation; protein folding/refolding; glycoprotein; glycosylation; protein structure

INTRODUCTION

Protein aggregation is a nonproductive consequence of protein folding. Moreover, protein aggregates have been linked to immunogenicity in animal models[1] and are implicated in immunogenicity for therapeutic proteins.[2–5] Partially unfolded proteins with nonnative structure have been identified as precursors to aggregation competent intermediates, leading to the irreversible formation of multimeric protein complexes.[6] Protein aggregates are either covalently (irreversible) or noncovalently (reversible) bound.[7] Noncovalent aggregates are primarily formed through weak interactions, typically as a result of normally buried hydrophobic residues becoming exposed to the aqueous environment through localized unfolding, or from electrostatic association of charged residues.[8,9] Noncovalent aggregation of proteins is reversible either through disruption of the polar environment with chaotropic agents (e.g., urea and GuHCl), surfactants that act to shield exposed hydrophobic patches, or by increasing the ionic strength of the buffering solution to disrupt electrostatic interactions. In contrast, covalent aggregates occur primarily as a result of nonnative, intermolecular disulfide-bond formation,[10,11] although other types of linkages are known.[12] Disulfide-linked protein aggregates tend to be dissociable only with the use of reducing agent. Additional aggregation classes have been identified, including conformationally altered and

Additional Supporting Information may be found in the online version of this article. Supporting Information

Correspondence to: Jin Xu (Telephone: +978-934-3673; Fax: +978-934-3013; E-mail: Jin_Xu@uml.edu)

Journal of Pharmaceutical Sciences, Vol. 102, 441–453 (2013)

chemically modified monomer species, nucleation-driven and surface-contact-driven aggregates.[13]

Aggregation of therapeutic proteins from recombinant cells that may occur during production is also a concern from a product-quality standpoint. Expression of proteins in prokaryotic systems yields high-protein titers with relatively low cost. However, in many cases, prokaryotic expression systems cannot produce properly folded, active protein from a eukaryotic genetic sequence, necessitating development of cumbersome protein-refolding protocols. Additionally, the lack of glycosylation machinery in prokaryotic expression systems results in proteins lacking the posttranslation modification. To resolve these issues, overexpression of protein therapeutics in Chinese hamster ovarian (CHO) cells has been widely utilized to create proteins with more eukaryotic characteristics including posttranslational modifications. As a result of higher levels of expression in the cell, the quality-control mechanisms of protein folding can become overwhelmed, resulting in production of partially processed and misfolded protein prone to aggregation.[14] Various other factors can also contribute to an increase in aggregation during therapeutic protein processing: storage time and temperature, protein concentration, ionic strength, agitation, posttranslational modification, formulation excipients, surface contacts, and so on.[7,15–18] All formulations, process materials, and procedures must be carefully chosen to minimize protein aggregation at all stages of manufacturing and processing.

As therapeutic protein aggregates are undesirable in the final purified product, downstream purification processes need to be optimized to remove the majority of protein aggregates that form during cell culturing. Multiple analytical methods are used for monitoring the formation of protein aggregates reliably and reproducibly [size-exclusion chromatography (SEC), sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), analytical ultracentrifugation, etc.]. Understanding the mechanisms of protein aggregation can allow aggregation during the therapeutic protein production process to be controlled. Significant efforts have been devoted to understanding the mechanisms of aggregation for immunoglobulin G (IgG) molecules that constitute the majority of currently approved protein therapeutics.[8,10,15,18–20] Fc-fusion proteins constitute a smaller proportion of approved therapeutics but are structurally similar to IgG in that the Fc portion of IgG is utilized as a scaffold for a receptor or protein ligand to improve its solubility, stability, and ease of purification. Investigations into Fc-fusions' aggregation mechanisms, as well as IgG's, have not been addressed, with only a few studies found in the literature.[21–24] Characterization of the soluble high-molecular-weight (HMW) aggregates produced during manufacturing of a highly disulfide-bonded and glycosylated Fc-fusion protein is presented here to better understand the potential aggregation mechanisms for this growing class of biotherapeutic proteins.

In our study, we used a model system Fc-fusion protein—activin receptor-like kinase 1 Fc (ALK1-Fc)—to study the aggregation mechanisms of its soluble HMW aggregates. The general structure of ALK1-Fc is shown in Figure 1a. ALK1 extracellular domain (ECD) is fused with the Fc region of human IgG and forms a covalently linked homodimer with another ALK1-Fc polypeptide chain via two disulfide linkages in the hinge region.[25] The "monomer" for the ALK1-Fc molecule is defined as the disulfide-bonded homodimer of the two main polypeptide chains. Including the two disulfides in the hinge region, the entire molecule has a total of 16 disulfide bonds—five in each ECD and two in each Fc. Figure 1b depicts the two expected products following IdeS protease treatment of ALK1-Fc: disulfide-bonded $(ECD)_2$ and a noncovalently associated Fc dimer. The potential *N*- and *O*-linked glycosylation sites for ALK1-Fc are outlined in Figure 1c. There are two potential *N*-linked sites

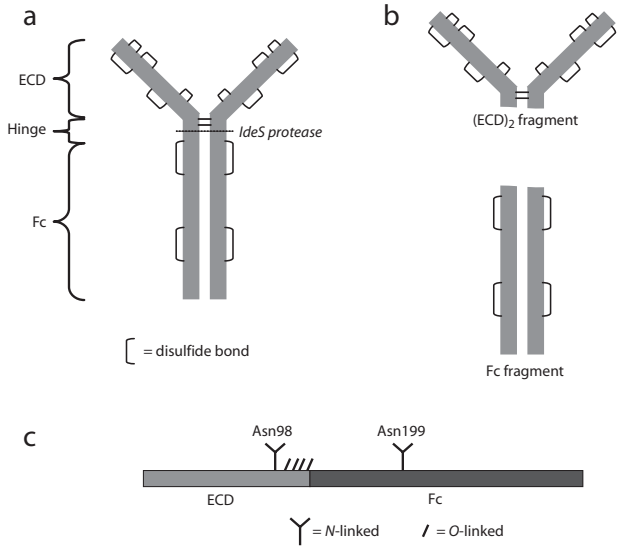

Figure 1. Cartoon diagram of ALK1-Fc fusion protein. (a) The "monomer" is defined as the homodimer of the two main chains, intermolecularly disulfide bonded in the hinge region. The entire molecule has 16 disulfide bonds in total. IdeS protease cleaves C terminal to the hinge-region intermolecular disulfide bonds (dotted line). (b) Proteolytic cleavage at this site yields two products: a bivalent (ECD)2 fragment and an Fc fragment. The two chains of the Fc fragment are held together noncovalently via electrostatic and hydrophobic interactions. (c) Two potential *N*-linked glycosylation sites are located at Asn98 and Asn199 in the ECD and Fc, respectively. Up to four potential *O*-linked glycosylation sites are possible in the hinge region.