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

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

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
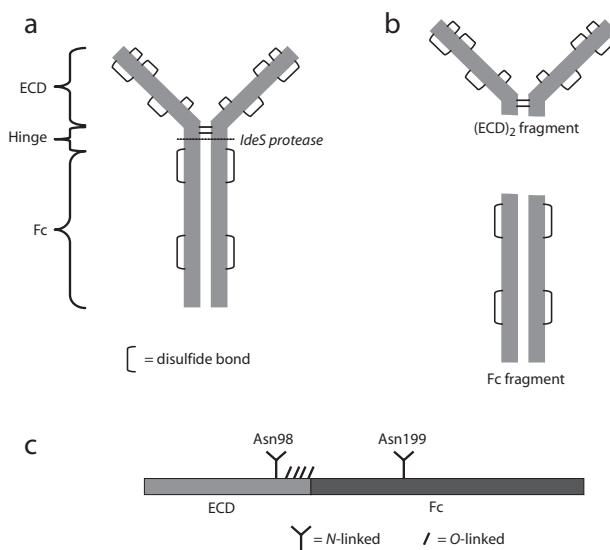

Figure 1. Cartoon diagram of ALK1-Fc fusion protein. (a) The "monomer" is defined as the homodimer of the two main chains, intermolecularly disulfide bonded in the hinge region. The entire molecule has 16 disulfide bonds in total. IdeS protease cleaves C terminal to the hinge-region intermolecular disulfide 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.

at Asn 98 and Asn 199 and up to four *O*-linked sites in the hinge region.

To better understand the aggregation mechanism of an Fc-fusion protein, various analytical methods were applied to characterize and analyze the soluble HMW aggregate species from the model system. Nonreducing SDS-PAGE analysis has shown that the majority of the aggregate species purified from an SEC column are covalently linked through nonnative, intermolecular disulfide bonds, with a small population associated through noncovalent interactions. Primary structural characterization by SEC–multiangle laser light scattering (MALLS), circular dichroism (CD) spectroscopy, and analysis of *N*- and *O*-linked glycosylation from peptide mapping revealed that the secondary structure and glycan microheterogeneity varies depending on the size of the aggregates. Intact masses of $(ECD)_2$ fragments generated from IdeS protease digestion of ALK1-Fc monomer and dimer were determined, and confirmed that intermolecular disulfide-bond formation of the dimer occurs through the ECD. Treatment with L-cysteine was sufficient to initiate covalent aggregation of purified monomer via disulfide-bond shuffling. Deglycosylation with PN-Gase F exacerbated the effect of treatment with cysteine. These studies provide a better understanding of Fc-fusion aggregation mechanisms and may allow for strategies limiting their formation during the manufacturing processes.

MATERIALS AND METHODS

Materials

Activin receptor-like kinase 1 Fc was expressed in CHO cells as previously described[25] and purified by recombinant protein A (proA) affinity chromatography. IdeS protease (Fabricator® IgG protease) was from Genovis and endoproteinase Glu-C was from Thermo Scientific.

Methods

Semipreparative SEC

Purified ALK1-Fc fusion protein was fractionated on a Sepax SRT-300 SEC column with an Agilent 1100 high-performance liquid chromatography (HPLC) system in phosphate-buffered saline (PBS) (pH 7.2) monitored by UV absorbance at 280 nm to separate HMW species from monomer. To facilitate downstream analysis, all fractions were further concentrated to 1 mg/mL with Amicon 10 K MWCO 4 mL centrifugal concentrators in a Beckman J-10 benchtop centrifuge.

Nonreducing and Reducing SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed using 4%–12% Bis–Tris/ MOPS NuPage gels (Invitrogen) at 200 V constant voltage for 50 min. Nonreduced samples were alkylated with 15 mM iodoacetic acid (IAA) and incubated for 10 min at 70°C prior to gel loading. Reduced samples were treated with 20 mM dithiothreitol (DTT) for 10 min at 70°C prior to gel loading. Detection of protein was accomplished with SimplyBlue SafeStain coomassie stain (Invitrogen) for 1 h following a 30-min wash with Milli-Q water. Identical amounts (1 μg) of each sample were loaded for normalization, while all electrophoretic mobility measurements and densitometric analyses were performed by ImageQuant software (GE Healthcare).

SEC–MALLS

Size-exclusion chromatography-fractionated ALK1-Fc was rechromatographed on a Sepax SRT-300 SEC column in PBS (pH 7.2) at 0.9 mL/min, prior to concentration, in-line with a Wyatt miniDAWN™ TREOS three-angle light scattering instrument.[26] An Optilab® T-rEX refractive index detector was used for concentration determination.

Peptide Map Analysis

Size-exclusion chromatography fractions 1–6 were reduced with 2.5 mM DTT in 6 M GuHCl/100 mM Tris–HCl (pH 7.5) for 1 h at 37°C and alkylated with IAA for 1 h at 25°C in the dark. An additional 2.5 mM DTT was added to quench any unreacted IAA. Reduced/alkylated protein was then desalted with 2 mL Zebaspin™ (Pierce) desalting columns into 50 mM Tris–HCl (pH 7.5) to remove any contaminating salts. Following desalting, 50 μg of protein was digested overnight at 37°C with a 1:20 (enzyme:substrate) ratio of endoproteinase Glu-C (Thermo Scientific). Peptides generated from overnight digestion were separated on a Phenomenex Jupiter 4 μm Proteo 90 Å C18 column (250×2 mm^2) at 1 mL/min in 0.05% trifluoroacetic acid (TFA) with a 0%–35% acetonitrile gradient delivered in 100 min by an Agilent 1200 HPLC instrument. Eluted peptides were detected by absorbance at 214 nm, with accurate mass determinations made by an LTQ Orbitrap mass spectrometer. Relative quantitation of *N*- and *O*-linked glycopeptide species was calculated from extracted ion chromatogram (XIC) peak areas normalized to the total area of XICs for all glycopeptide species searched (e.g., Man5 and bi-, tri-, and tetra-antennary for *N*-linked glycopeptides and unoccupied and singly, doubly, and triply occupied *O*-linked glycopeptide species). Mass determinations for the glycans were calculated from the accurate mass of the entire glycopeptide minus the molecular weight of the peptide sequence. Determination of *N*-linked glycan structures was performed using a combination of the GlycoMod tool[27] and known eukaryotic oligosaccharide biosynthetic pathways.[28]

CD Spectroscopy

Activin receptor-like kinase 1 Fc protein was diluted to 0.2 mg/mL in 10 mM sodium phosphate (pH 7.8) in a 0.1-cm path-length quartz cuvette. CD-spectroscopy measurements were taken on a Jasco J-810 CD spectrapolarimeter in the far-UV range from 250 to 190 nm at 20°C. All spectra were subtracted against buffer blanks to eliminate any contribution from the buffer.[29]

Middle–Down Analysis by Mass Spectrometry

Control protein (100 μg), disulfide-linked trimer/tetramer, and disulfide-linked dimer were each treated with PNGase F and sialidase A to remove all N-linked oligosaccharides and sialic acids prior to digestion for 3 h at 37°C with 10 μg of IdeS protease.[30] Separate digestions were conducted with the monomer, dimer, and trimer/tetramer in native form (i.e., fully glycosylated) for analysis by nonreducing SDS-PAGE. The deglycosylated digests were eluted from an Xbridge (Waters) C4 reversed-phase liquid chromatography column on an Agilent 1200 HPLC system to separate the extracellular domain from the Fc, with in-line mass detection by an LTQ Orbitrap mass spectrometer. Mass spectral deconvolution of fragment charge envelopes was attained with the ProMass deconvolution software package (Thermo Scientific).

In vitro Aggregate Generation

Purified ALK1-Fc monomer (native and de-N-glycosylated/desialylated) at 1 mg/mL in PBS (pH 7.2) was treated with 2 mM L-cysteine at 37°C for 0–8 h. Aliquots of 5 μg were removed at the times indicated, SDS sample loading buffer was added with 15 mM IAA, and finally frozen at −80°C. For final analysis, the frozen aliquots were analyzed by nonreducing SDS-PAGE as outlined above.

RESULTS

Purification of Soluble HMW Aggregate Species and Size Analysis

Activin receptor-like kinase 1 Fc fusion protein was first isolated from clarified cell culture medium by a single-step proA affinity purification. To analyze the HMW species of ALK1-Fc fusion protein, six fractions were isolated from multiple injections on an analytical-scale SEC column, spanning 5.5–8.5 min of the A_{280} chromatogram (Fig. 2a), ∼30 s per fraction. On the basis of total peak areas of the HMW and monomer fractions of the SEC chromatogram, ALK1-Fc fusion protein eluted from the proA column is composed of ∼32% HMW species and ∼68% monomer. This result is in good agreement with orthogonal analysis by analytical ultracentrifugation

Table 1. Molecular Weight of SEC Fractions by MALLS

Fraction	Retention time (min)	MW (kDa)	Polydispersity
1a	6.11	–	1.027 (6%)
1b	8.84	–	1.605 (94%)
2a	6.18	–	1.040 (4%)
2b	6.55	814	1.038 (7%)
3	7.07	372	1.006 (9%)
4	7.43	276	1.043 (3%)
5	7.95	192	1.005 (4%)
6	8.69	94	1.009 (7%)

(AUC), which reported ∼34% HMW species. The fractions are herein referred to as Frac-1 for SEC fraction one, Frac-2 for SEC fraction two, and so on. Frac-1–Frac5 were collected in the region of the chromatogram corresponding to HMW aggregate species. Frac-6 was collected in the region corresponding to monomer.

Rechromatography of the SEC fractions with MALLS detection (Fig. 2b) revealed that Frac-1 partially rechromatographs as HMW aggregate (peak a in Fig. 2b) and also as monomer (peak b in Fig. 2b). Frac-2–Frac-6 eluted from the SEC column with retention times consistent with the initial SEC elution used to generate the fractions (Table 1). The molecular sizes of Frac-1's two peaks were not determined from the MALLS data because of the lack of a monodisperse sample; the same observation was made for Frac-2, peak a. Peak b for Frac-2 was measured by MALLS at 814 kDa (between octamer and nonamer); however, this size is approaching the upper limit of accuracy for the three-angle light-scattering instrument and may not accurately represent the molecular weight of the aggregate species in Frac-2. Frac-3–Frac-5 have low polydispersity and have molecular weight values of 372, 276, and 192 kDa, respectively. Using the molecular weight calculated for the monomer from the MALLS data for Frac-6 (94 kDa), it was determined that Frac-3–Frac-5 represent tetramer, trimer, and dimer, respectively.

Reducing and nonreducing SDS-PAGE of the concentrated SEC fractions (Fig. 3) indicated that the majority of the HMW aggregates observed are covalently linked through disulfide bonding and are not dissociated by SDS denaturation alone. When Frac-2–Frac-5 were treated with reducing agent, the primary aggregate forms (>98% band-area fraction) were reduced to sizes consistent with reduced monomer (SEC Frac-6). The sizes of the disulfide-linked aggregates predictably decrease with increasing retention time from the SEC column. In agreement with the MALLS data, nonreducing SDS-PAGE of SEC Frac-5 gives a band consistent with a dimer of molecular weight ∼175 kDa, on the basis of electrophoretic mobility. Frac-4 is primarily composed of trimer and lower amounts of tetramer. Frac-3 has two
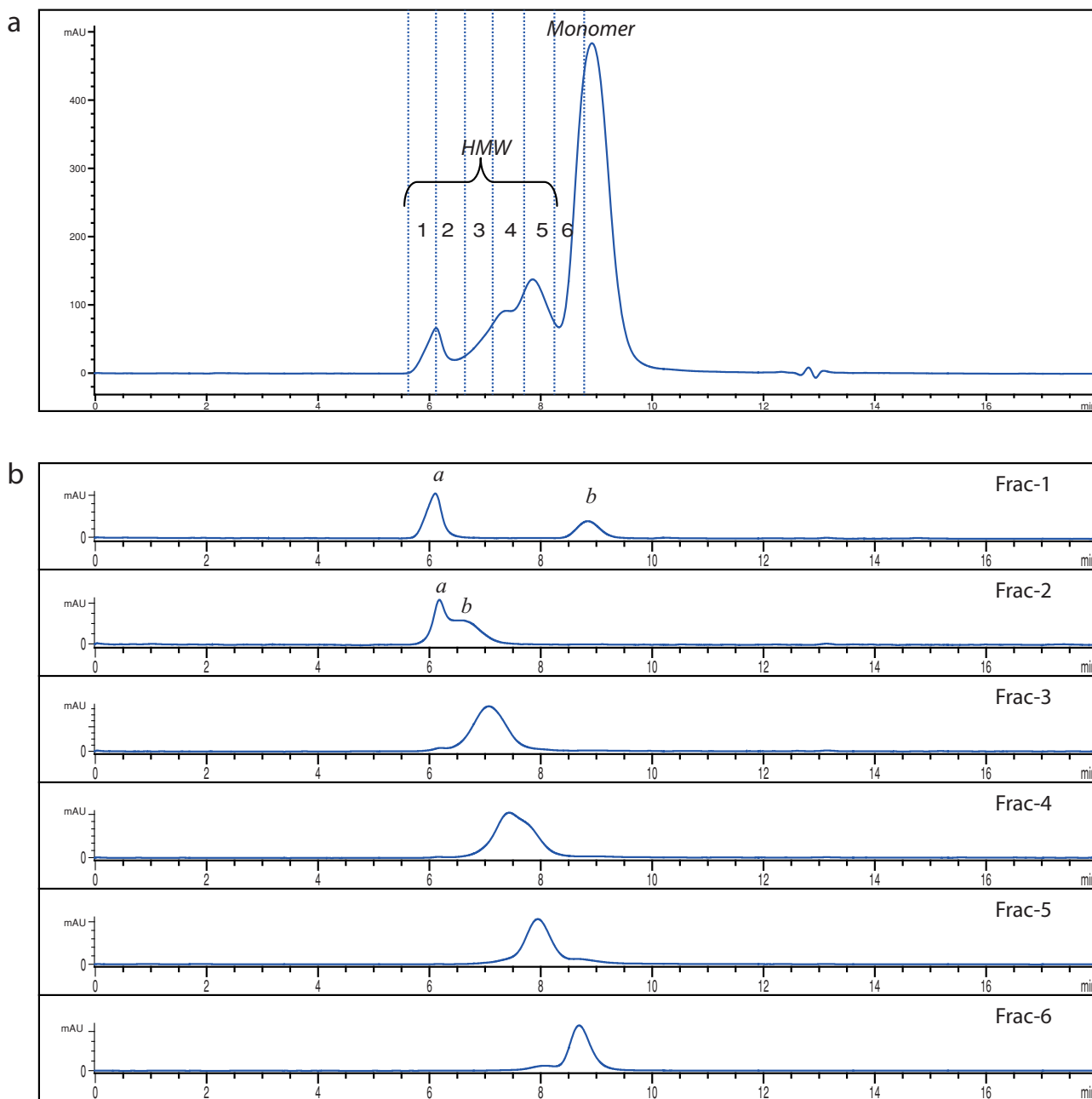
Figure 2. Semipreparative SEC of ALK1-Fc fusion protein. (a) Protein A purified ALK1-Fc fusion protein eluted from an analytical-scale SEC column. Vertical dashed lines represent the regions used for fractionation. (b) Rechromatograph of semipreparative SEC fractions in-line with MALLS shows that Frac-1 is partially converted to monomer as a result of dilution. Frac-2–Frac-5 rechromatograph in the same positions as the semipreparative run used for fractionation. Frac-6 rechromatographs with a retention time similar to the monomer peak.

bands representing mostly tetramer and a smaller amount of trimer. The various aggregate species do not completely separate by SEC, leading to considerable overlap between the fractions. Only the dimer fraction (Frac-5) elutes from an SEC column with a defined peak (52% of all HMW species by SEC; 54% by AUC). Frac-2 does not have a discernible band on the nonreducing gel because a large majority of the protein aggregates do not enter the gel. The gel shows that a population of HMW aggregates from Frac-1 is formed through noncovalent interactions that are dissociable by SDS treatment. The size of the monomer generated from SDS dissociation for the noncovalent aggregate found here is roughly 93 kDa by nonreducing SDS-PAGE analysis, comparable in size to the native monomer for ALK1-Fc (94 kDa by SEC–MALLS).
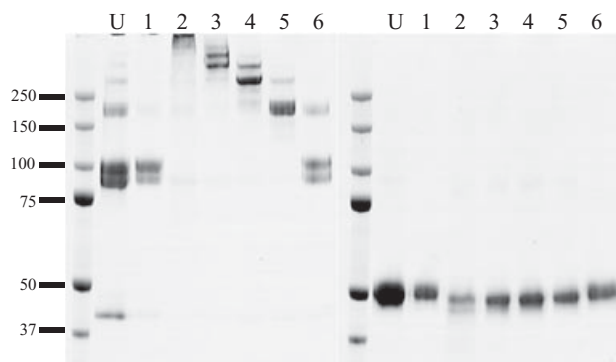
Figure 3. SEC fractions by SDS-PAGE. Nonreducing (left) and reducing SDS-PAGE (right) for ALK1-Fc SEC fractions. Unfractionated proA elution is included as a control (U). Frac-1 is a mixture of noncovalent and covalent aggregates, whereas Frac-2–Frac-5 are all covalently linked through intermolecular disulfide bonds. Frac-6 is primarily monomer, with a small amount of covalent dimer.

An additional SEC analysis of Frac-1 (Fig. S1) at higher ionic strength (500 mM NaCl) further demonstrates that Frac-1 contains salt-dissociable, noncovalent aggregates (Fig. S1). Additionally, Frac-1 has a second population of HMW species that are large, disulfide-linked aggregates that do not enter the gel, similar to Frac-2.

Primary Structural Analysis of HMW Aggregates Species

Characterization of the *N*- and *O*-linked glycan moieties of all of the HMW species isolated from the SEC elution was carried out using a Glu-C endoproteinase peptide map procedure. Two of the peptides produced from Glu-C endoproteinase digestion of ALK1-Fc are found with *N*- and *O*-linked posttranslational modifications present, representing the ECD *N*-linked site and potential *O*-linked glycosylation sites near the hinge region. The resultant combined mass spectra spanning the entire elution of the *N*-linked glycopeptide for SEC Frac-3, Frac-5, and Frac-6 are shown in Figure 4. The *N*-linked glycopeptide is occupied by mostly bi- and triantennary complex glycans with varying levels of sialylation, as well as some oligomannose moieties, particularly Man5 (1216.42 Da in Fig. 4). The higher-order HMW species have increased relative levels of Man5 compared with the lower-order HMW species (*e.g.*, dimer) and ALK1-Fc monomer. On the basis of peak areas from the XIC of the Man5 glycopeptide normalized to the total peak area contribution of all *N*-linked glycopeptide species, we see a decrease in the Man5 glycopeptide from Frac-3 to Frac-6. Conversely, an increase in tri- and tetra-antennary, complex glycopeptides is observed when comparing Frac-3–Frac-6. On the basis of the peptide map data, the HMW species isolated from the SEC elution tend to

have less-processed, less-sialylated *N*-linked glycan species at this site of the ECD.

Analysis of the *O*-linked glycosylation occupancy was performed in a similar manner as the ECD *N*-linked glycosylation site, utilizing a hinge-region glycopeptide from the endoproteinase Glu-C peptide map known to have up to four *O*-linked glycans. The *O*-linked glycan moieties in this region are known to be Core 1, mucin-type structures. The resultant combined mass spectra for the hinge-region *O*-linked glycopeptide are shown in Figure 5. On the basis of peak areas from the XICs of unoccupied peptide and singly, doubly, and triply occupied peptide normalized to the XICs of all *O*-linked glycopeptides in this region, the levels of unoccupied peptide decrease steadily as the size of the HMW species becomes smaller. The higher-order HMW species from Frac-1 and Frac-2 have a higher relative level of unoccupied *O*-linked glycans compared with Frac-3–Frac-5 and the monomer from Frac-6. There is an increase in glycopeptides with three *O*-linked glycans in the hinge region, when comparing monomer with all of the aggregate species. On the basis of these observations, the *O*-linked glycan species for the HMW aggregates isolated from the SEC column are also less processed and less sialylated compared with the monomer.

Higher-Order Structural Analysis of HMW Aggregate Species

Circular dichroism spectroscopy was utilized to analyze and compare the secondary structure and folding of the HMW aggregate species with monomeric ALK1-Fc. Analysis of each of the SEC fractions on a CD spectropolarimeter in the far-UV range (250–190 nm) showed that Frac-4–Frac-6 (Fig. 6a) have secondary structure characteristics that closely resemble the purified ALK1-Fc control. Two minima are evident from the data for each of the spectra: one at 195 nm and the other at 213 nm. The 195-nm minimum is usually associated with random-coil secondary structure. The 213-nm minimum observed is very close to the typical minimum of 218 nm normally attributed to the β-sheet secondary structure found in IgG.[28] Separate spectra collected for ALK1 ECD and IgG Fc (data not shown) indicate a strong minimum at 201 nm for the ECD that distorts the minimum at 218 nm for the β-sheet secondary structure of the Fc region, resulting in the minimum of 213 nm observed for ALK1-Fc. A single maximum is visible for the purified control ALK1-Fc at 203 nm, a known maximum for IgG and other proteins with high β-sheet content. For Frac-4 and Frac-5, the maximum in this region appears to shift closer to 200 nm as the size of the HMW species increases. Overall, the spectral differences between monomeric ALK1-Fc (from the purified control and Frac-6) and SEC Frac-4 and Frac-5 are small, indicating that the lower-order HMW aggregates (trimer
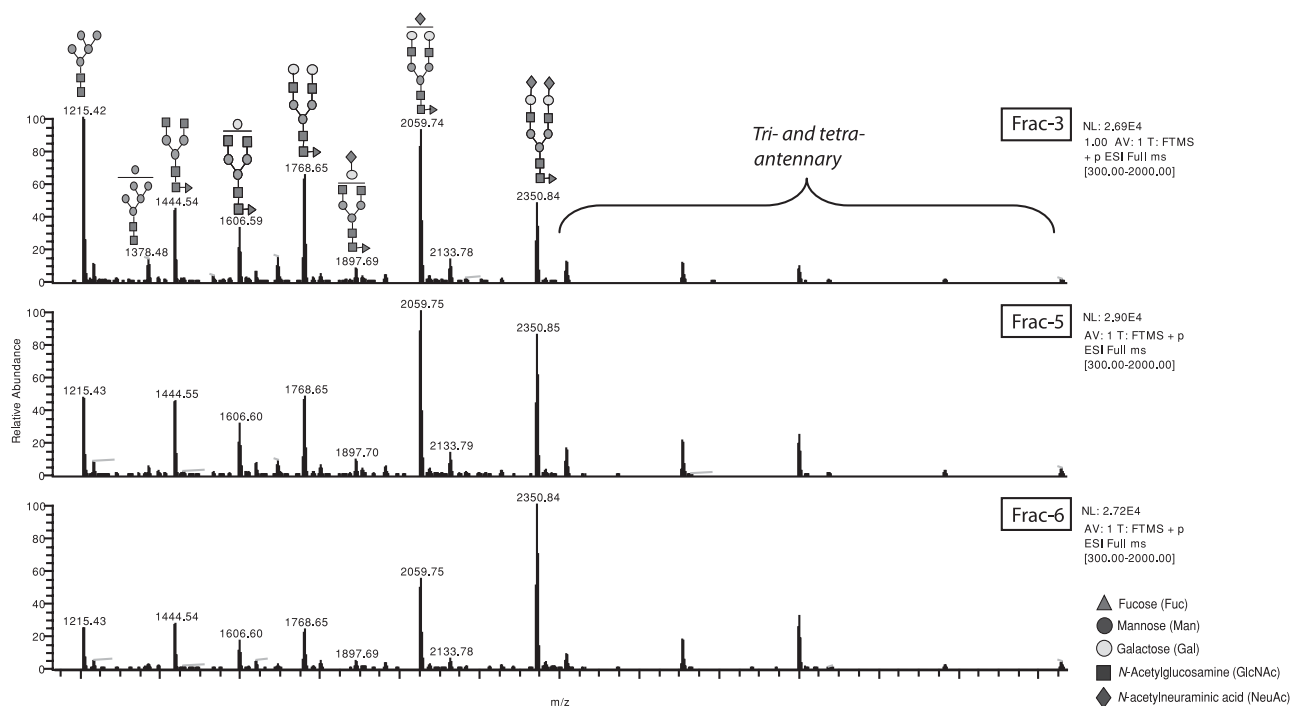
Figure 4. *N*-linked glycan analysis of SEC fractions by peptide mapping. Reduced and alky-lated ALK1-Fc SEC fractions were digested with Glu-C endoproteinase. Masses shown have been subtracted from the mass of the unoccupied peptide and reflect only the mass of the oligosaccharides. The combined mass spectra show that the site is primarily occupied by bi-, tri-, and tetra-antennary complex glycans. Relative quantification of species was based on nor-malized peak areas from the extracted ion chromatograms. Higher-order aggregates have higher relative levels of oligomannose and neutral species compared with smaller oligomers.

and dimer) maintain overall native secondary struc-ture and folding, comparable with the purified control ALK1-Fc and Frac-6.

As the HMW species from the SEC fractionation increase in size, larger spectral differences are ob-served by CD. Frac-1–Frac-3 (Fig. 6b) are composed

of the largest HMW species from the SEC elution. Frac-1 and Frac-2 CD spectra (open circles and open triangles, respectively, in Fig. 6b) are similar. Both have strong negative responses with minima at 210 and 215 nm. At 210 nm, there is a strong rise in the signal that reaches positive values at 200 nm and
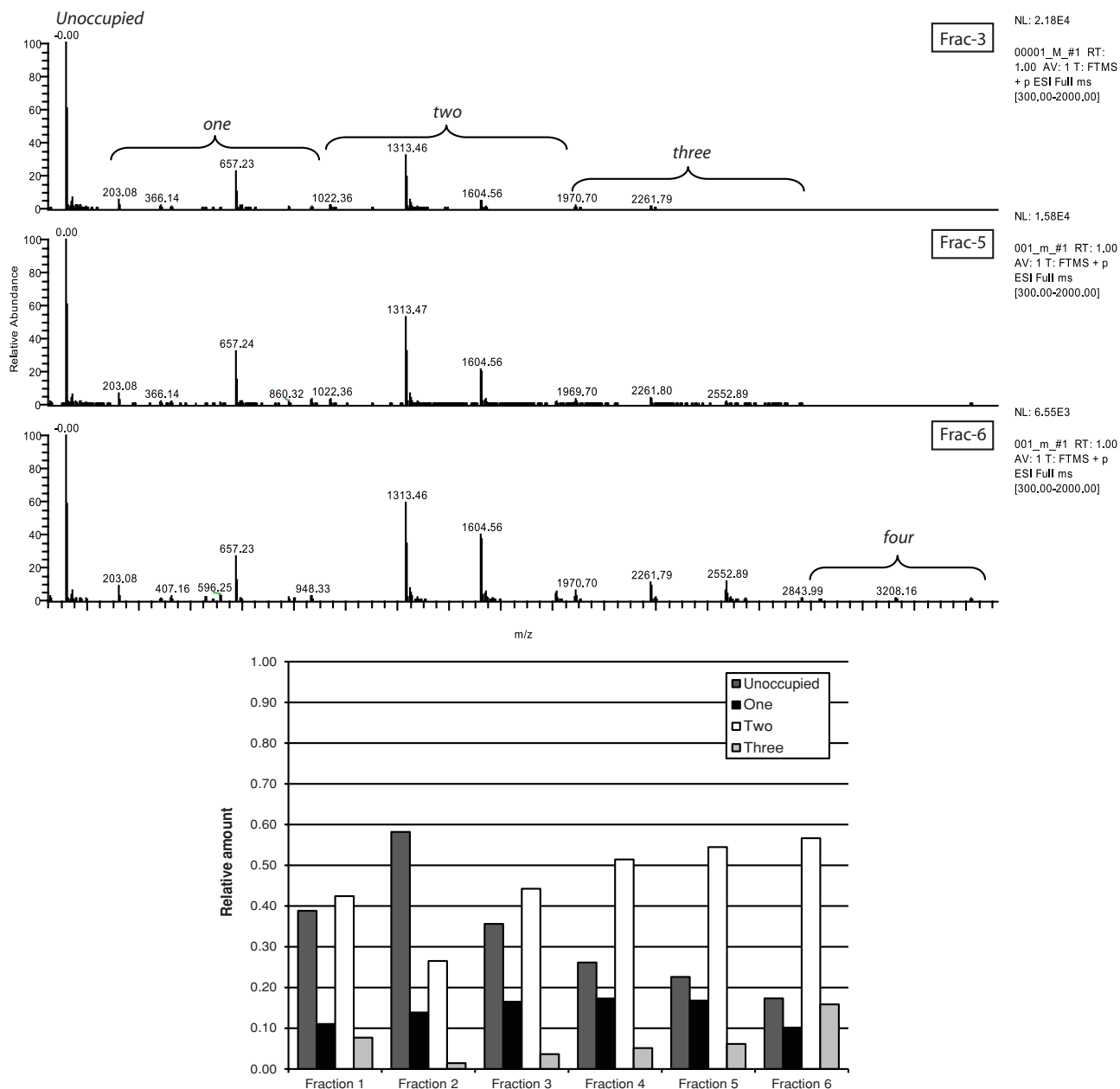
Figure 5. Combined mass spectra from peptide mapping analysis covering *O*-linked occupancy in the hinge region. *O*-linked occupancy is labeled with braces. Unoccupied species labeled accordingly. Masses shown have been subtracted from the mass of the unoccupied peptide and reflect only the mass of the oligosaccharides. Higher-order aggregate fractions (e.g., Frac-3) have less occupancy overall when compared with monomeric ALK1-Fc. Relative quantification of species was based on normalized peak areas from the extracted ion chromatograms.

continues to 9000 deg cm^2/dmol at 190 nm. The maximum of the signal in the lower wavelength region cannot be determined for Frac-1 or Frac-2 because these spectra never reach a peak before the lower wavelength cutoff of 190 nm. Spectra obtained below 190 nm were very noisy using phosphate buffer and are not reported here. The differences in the spectra of Frac-1 and Frac-2 compared with the monomeric ALK1-Fc indicate that the higher-order aggregate species found in these fractions have significant secondary structure differences as compared with the monomer or the lower-order aggregate species (e.g., dimer, trimer, etc.), suggesting differential folding of the higher-order aggregates. Frac-3 CD spectra (open diamonds in Fig. 6b) appear to be an intermediate fold, having spectral qualities that resemble both the properly folded ALK1-Fc control and the differently folded, high-order HMW species of Frac-1 and Frac-2.
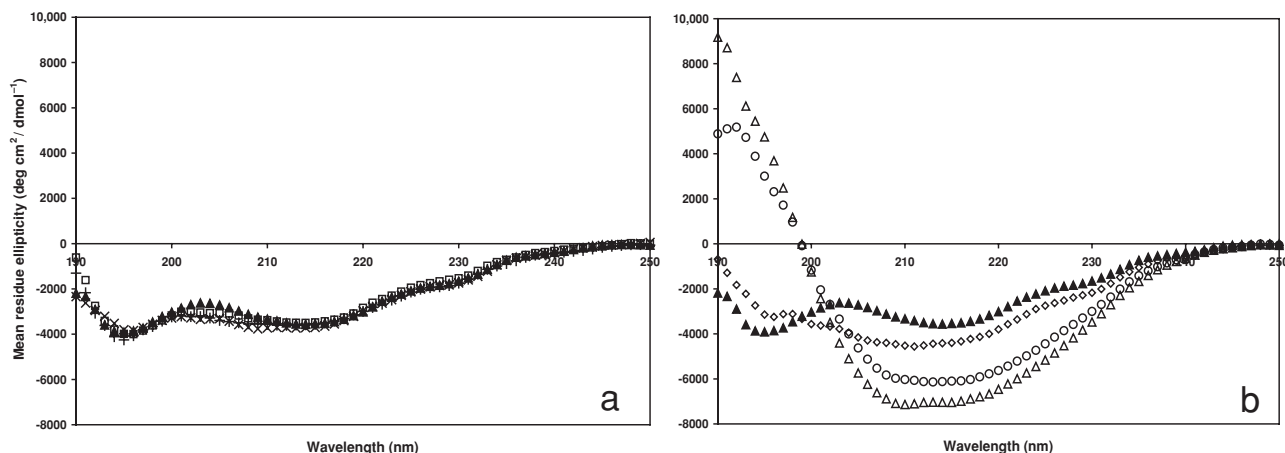
Figure 6. CD spectroscopy of SEC Frac-1–Frac-6 with control ALK1-Fc. CD spectra plotted as a function of mean residue ellipticity and wavelength in the far-UV region from 250 to 190 nm. The control spectrum for ALK1-Fc is used in each panel for reference. (a) Control (▲), Frac-6 (□), Frac-5 (+), and Frac-4 (×). Similarity in spectra for dimer and trimer HMW species in Frac-5 and Frac-4, respectively, indicates that the secondary structure for lower-order, covalent aggregates is relatively unaffected. (b) Control (▲), Frac-3 (◇), Frac-2 (△), and Frac-1 (○). As the HMW species become larger (tetramer and larger), a significant deviation in folding from the control spectra is evident.

Covalent HMW Aggregates of ALK1-Fc are Disulfide Linked Through the ECD

On the basis of the distribution of HMW species from the SEC elution, an additional SEC column fractionation was conducted to isolate disulfide-linked dimer, trimer, and tetramer. Two fractions corresponding to dimer (lane D1 in Fig. 7a) and a mixture of trimer and tetramer (lane T1 in Fig. 7a) were isolated for comparison with a monomer control for ALK1-Fc (lane M1 in Fig. 7a). Monomer, dimer, and trimer/tetramer fractions were treated with IdeS protease, generating an $(ECD)_2$ and a noncovalently linked Fc (Fig. 1b). The resultant products of the digestion were analyzed by nonreducing SDS-PAGE (Fig. 7a). The digest of the monomer fraction (lane M2 in Fig. 7a) shows two distinct bands corresponding to the $(ECD)_2$ product (45 kDa) and Fc portions (26 kDa) of ALK1-Fc from the digest. The dimer and trimer/tetramer digests (lanes D2 and T2, respectively, in Fig. 7a) have a single distinct band corresponding to the Fc portion of ALK1-Fc, and a second region of staining with molecular weights between 70 and 100 kDa for the dimer fraction and 100–250 kDa for the trimer/tetramer fraction. The 70–100 kDa region of staining for the dimer fraction presumably corresponds to the $(ECD)_2$ portion of ALK1-Fc, intermolecularly disulfide bonded to a second $(ECD)_2$. The staining of the ECD aggregates is poor because of heterogeneous *N*- and *O*-linked glycosylation and the resulting diffuse banding pattern. The presence of the Fc band at 25 kDa and the absence of $(ECD)_2$ in the dimer and trimer/tetramer digests suggest that the covalent HMW aggregate species were formed through the ECD portion of the molecule.

To confirm that the covalent HMW aggregates are indeed formed through disulfide bonds in the ECD portion of ALK1-Fc, all digested fractions were also analyzed by C4 RP-HPLC in-line with an Orbitrap mass spectrometer for intact mass analysis. All digest products were first treated with sialidase A and PNGase F before injection onto the C4 column to eliminate glycan microheterogeneity that could complicate the intact mass analysis. The A_{280} chromatograms for the C4 elutions of the monomer, dimer, and trimer/tetramer fractions (Fig. 7a) display two peaks. All three chromatograms have a peak at retention time (RT) = 14.8 min with a mass of 23,815 Da that is consistent with the theoretical mass of the Fc-region digestion product generated from IdeS protease treatment. The monomer digest has a second peak at RT = 11.4 min, whose intact mass matches the expected molecular weight of the $(ECD)_2$ generated from proteolytic treatment (27,344 Da). The second peak of the dimer digest with RT = 11.7 min also corresponds to the ECD portion of ALK1-Fc but with a mass consistent with a covalent dimer of $(ECD)_2$ (54,690 Da). Unfortunately, the trimeric $(ECD)_2$ peak did not produce deconvolutable mass spectra. On the basis of the intact mass data, it is likely that the disulfide-linked aggregate species for ALK1-Fc are formed through the ECD and not the Fc region for the dimer, with the other, higher-order HMW species likely formed in the same manner.
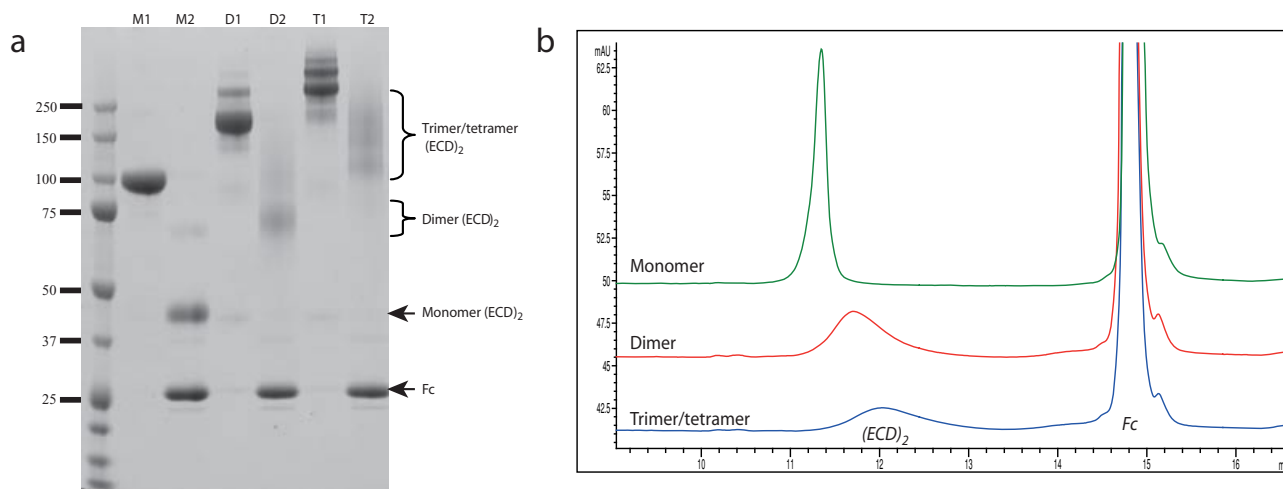
Figure 7. IdeS protease treated ALK1-Fc monomer, dimer, trimer/tetramer. (a) Nonreducing SDS-PAGE of untreated (1) and digested (2) ALK1-Fc. Digestion of monomer (M2) results in two bands with molecular weights consistent with (ECD)2 (~44 kDa) and Fc (~26 kDa) portions of the molecule. Digestion of dimer and trimer/tetramer fractions (D2 and T2, respectively) also produces Fc band, with (ECD)$_2$ dimer identified from 70 to 100 kDa for the dimer fraction and (ECD)2 oligomers between 100 and 250 kDa for the trimer/tetramer fraction. (b) A_{280} chromatogram for C4 reversed-phase elution of digested fractions: monomer (green), dimer (red), and trimer/tetramer (blue).

Mechanism of ALK1-Fc Aggregation

Because covalent HMW aggregates of ALK1-Fc form through intermolecular disulfide bonding, the effect of introducing a free sulfhydryl (in the form of L-cysteine) was tested. An effective concentration of L-cysteine was determined empirically from titration experiments ranging from 0.25 to 50 mM L-cysteine incubated with ALK1-Fc at a concentration of 1 mg/mL in PBS (pH 7.2) at 37°C. A concentration of 2 mM L-cysteine was chosen for the final time-course experiment, the results of which were analyzed by nonreducing SDS-PAGE (Fig. 8).

Glycosylation is thought to have a protective effect on protein aggregation, and the impact of glycosylation on ALK1-Fc aggregation was evaluated. Both native and sialidase A/PNGase F-treated (deglycosylated) ALK1-Fc were evaluated for 8 h at 37°C in the presence of 2 mM L-cysteine. Nonreducing SDS-PAGE analysis of the native protein shows increased dimer formation after only 1 h of incubation at 37°C for both native and deglycosylated protein. However, deglycosylated ALK1-Fc displayed a much higher level of HMW aggregate formation than the native form (64% vs. 20% HMW species) following 8 h of incubation at 37°C. The deglycosylated ALK1-Fc has formed HMW species consisting of covalently linked dimer, trimer, tetramer, pentamer, hexamer, and septamer; whereas no HMW species above trimer were observed for the native protein by nonreducing SDS-PAGE. Continuing the incubation past 8 h resulted in the formation of a visible precipitate after 16 h for
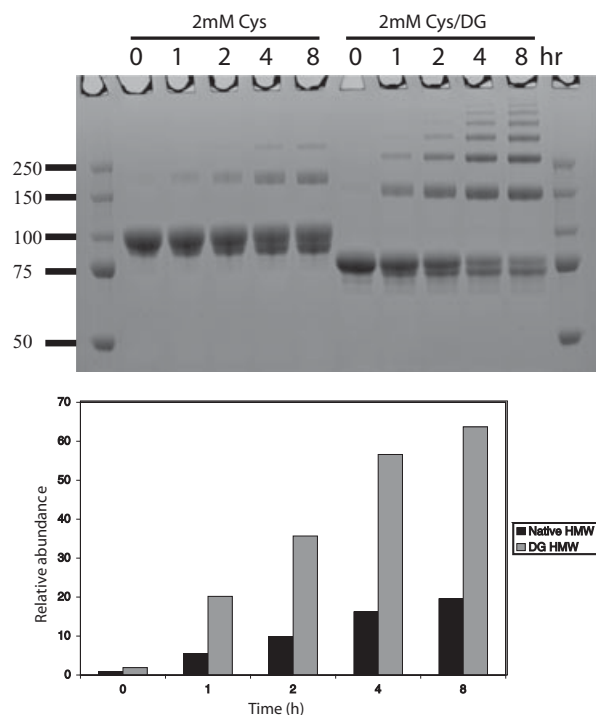

Figure 8. Native and deglycosylated (DG) ALK1-Fc with 2 mM L-cysteine. Nonreducing SDS-PAGE of 8-h time course of ALK1-Fc incubated at 37°C. Deglycosylated ALK1-Fc (DG, desialylated, and deN-glycosylated) shows increased aggregation propensity as a result of disulfide-bond shuffling. Relative amounts of HMW species by densitometry (dimer g octamer) are compared for native and DG ALK1-Fc. After 8 h, DG ALK1-Fc is composed of ~65% HMW species, whereas native is ~20%.

both the native and deglycosylated samples (data not shown).

DISCUSSION

The soluble HMW aggregate species from ALK1-Fc fusion protein following initial proA purification have been characterized, with the intention of dissecting the mechanism of their formation. The HMW species formed are primarily irreversible intermolecularly disulfide-bonded oligomers. Other types of covalent linkages may be involved, on the basis of the nonreducible species observed as minor species in the reducing SDS-PAGE analysis, but are outside the scope of this paper. A small population of aggregates from Frac-1 is reversible and may form through noncovalent bonds such as electrostatic interactions. It was possible to dissociate these noncovalent aggregate species using higher ionic strength during SEC (Fig. S1). These reversible aggregates are present in lower abundance than the covalent form and are typically removed through purification—therefore, the characterization presented here focused primarily on the covalent species. Multiple, orthogonal methods were employed in this work to characterize the covalent HMW species from the SEC fractionation. The results show that the aggregate is composed of various-size oligomers ranging from mostly dimer (∼50% of the HMW species) up to nonamer as determined by SEC–MALLS and AUC, with the likelihood that even larger oligomeric complexes are formed that are beyond the limits of available size measurement.

From the CD spectroscopy data, the highest molecular weight aggregate species from Frac-1 and Frac-2 have the largest structural differences when compared with the monomeric form of the protein found in Frac-6 and the purified control ALK1-Fc. Differential folding in Frac-1 may expose hydrophobic patches on the surface of the protein prone to noncovalent aggregation, helping explain the monomeric peak observed when rechromatographed on SEC at a lower protein concentration or when analyzed with SDS in its nonreduced form. The lower-order HMW species found in Frac-4 (trimer) and Frac-5 (dimer) have CD spectra that are largely dominated by the Fc portion of the molecule, with few noticeable changes in secondary structure compared with the monomeric form of the protein.

The HMW species were also observed to contain a reduced level of glycosylation in comparison with monomeric ALK1-Fc. De-N-glycosylated, monomeric ALK1-Fc was more susceptible to aggregation in the presence of a free sulfhydryl. This observation suggests that the less complex and less-extensively glycosylated ALK1-Fc in culture may be more susceptible to covalent HMW aggregate formation. PNGase F treatment of proteins has been shown to decrease a glycoprotein's thermal stability and to increase their propensity to self-associate, supporting the notion that the glycans on ALK1-Fc play a protective role.[31] Glycans are also known to stabilize protein structures, supporting the hypothesis that removal of glycan from ALK1-Fc may destabilize the native structure and favor an aggregation pathway.[32,33] Furthermore, it cannot be excluded that the native fold of the ECD may have been altered simply through deglycosylation because PNGase F treatment results in the formation of a nonnative aspartate in the place of asparagine. Although the complete removal of the N-linked glycans for the L-cysteine-treated ALK1-Fc is an extreme example, it does highlight the importance of glycan occupancy for prevention of aggregation and may further explain why the ALK1-Fc HMW aggregates are enriched with underprocessed forms of N- and O-linked glycans.

Current mechanisms of protein aggregation focus primarily on localized unfolding of the native structure as a precursor to an aggregation-competent intermediate.[16,34,35] During oxidative protein folding within the endoplasmic reticulum, protein disulfide isomerase (PDI) catalyzes disulfide-bond formation by oxidation of free sulfhydryl side chains of cysteine. In addition to promoting disulfide-bond formation, PDI can also reshuffle incorrectly formed disulfide bonds to their native state. The process is repeated a number of times until only native disulfide bonding and folding are attained.[36] Under conditions of cellular stress (i.e., overexpression, high temperature, etc.), incomplete oxidative protein folding may result in the secretion of small amounts of protein with unformed disulfide bonds into the culture medium. It has been shown that IgG expression in CHO cells results in a small fraction of protein secreted into the culture medium with unformed disulfide bonds (∼0.02–0.1 mol sulfhydryl per mol of protein).[37] We speculate that a similar mechanism may function for ALK1-Fc, in which a partially reduced intermediate species is secreted into the culture medium, where it can catalyze disulfide-bond shuffling to form the covalently linked HMW aggregate species observed during cell culturing.

This work has led to increased understanding of the aggregation mechanism for a model Fc-fusion protein expressed in CHO cells. Analysis of the specific domains implicated in covalent-aggregate formation for an IgG1 has been conducted.[38] To the best knowledge of the authors of this work, analysis of the domains involved in covalent-aggregate formation of Fc-fusion proteins has never been investigated. It is important to determine the mechanism of aggregation of this growing class of biotherapeutics so that conditions to minimize their formation can be implemented during production.

ACKNOWLEDGMENTS

REFERENCES

1. Braun A, Kwee L, Labow MA, Alsenz J. 1997. Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-α) in normal and transgenic mice. Pharm Res 14(10):1472–1478.

2. Carpenter JF, Randolph TW, Jiskoot W, Crommelin DJA, Middaugh CR, Winter G, Fan YX, Kirshner S, Verthelyi D, Kozlowski S, Clouse KA, Swann PG, Rosenberg A, Cherney B. 2009. Overlooking subvisible particles in therapeutic protein products: Gaps that may compromise product quality. J Pharm Sci 98(4):1201–1205.

3. Hermeling S, Crommelin DJ, Schellekens H, Jiskoot W. 2004. Structure–immunogenicity relationships of therapeutic proteins. Pharm Res 21(6):897–903.

4. Rosenberg AS. 2005. Effects of protein aggregates: An immunologic perspective. AAPS J 8(3):E501–E507.

5. Singh SK. 2010. Impact of product-related factors on immunogenicity of biotherapeutics. J Pharm Sci 100(2):354–387.

6. Kim YS, Randolph TW, Stevens FJ, Carpenter JF. 2002. Kinetics and energetics of assembly, nucleation, and growth of aggregates and fibrils for an amyloidogenic protein. J Biol Chem 30(277):27240–27246.

7. Cromwell MEM, Hilario E, Jacobsen F. 2006. Protein aggregation and bioprocessing. AAPS J 8(3):E572–E579.

8. Moore JMR, Patapoff TW, Cromwell MEM. 1999. Kinetics and thermodynamics of dimer formation and dissociation for a recombinant humanized monoclonal antibody to vascular endothelial growth factor. Biochemistry 38:13960–13967.

9. Lu HS, Chang WC, Mendiaz EA, Mann MB, Langley KE, Hsu YR. 1995. Spontaneous dissociation–association of monomers of the human-stem-cell-factor dimer. Biochem J 305(2):563–568.

10. Andya JD, Hsu CC, Shire SJ. 2003. Mechanisms of aggregate formation and carbohydrate excipient stabilization of lyophilized humanized monoclonal antibody formulations. AAPS Pharm Sci 5(2):E10.

11. Constantino HR, Langer R, Klibanov AM. 1994. Solid-phase aggregation of proteins under pharmaceutically relevant conditions. J Pharm Sci 83(12):1662–1669.

12. Ruf RAS, Lutz EA, Zigoneanu IG, Pielak GJ. 2008. Alpha-synuclein conformation affects its tyrosine-dependent oxidative aggregation. Biochemistry 47(51):13604–13609.

13. Philo JS, Arakawa T. 2009. Mechanisms of protein aggregation. Curr Pharm Biotech 10:348–351.

14. Schroder M, Schafer R, Friedl P. 2002. Induction of protein aggregation in an early secretory compartment by elevation of expression level. Biotech Bioeng 78(2):131–140.

15. Kueltzo LA, Wang W, Randolph TW, Carpenter JF. 2008. Effects of solution conditions, processing parameters, and container materials on aggregation of a monoclonal antibody during freeze-thawing. J Pharm Sci 97(5):1801–1812.

16. Chi EY, Sampathkumar K, Randolph TW, Carpenter JF. 2003. Physical stability of proteins in aqueous solution: Mechanisms and driving forces in non-native protein aggregation. Pharm Res 20(9):1325–1336.

17. Truehit MJ, Kosky AA, Brems DN. 2002. Inverse relationship of protein concentration and aggregation. Pharm Res 19(4):511–516.

18. Sahin E, Grillo AO, Perkins MD, Roberts CJ. 2010. Comparative effects of pH and ionic strength on protein–protein interactions, unfolding, and aggregation of IgG1 antibodies. J Pharm Sci 99(12):4830–4848.

19. Qian Y, Jing Y, Zheng JL. 2010. Glucocorticoid receptor-mediated reduction of IgG-fusion protein aggregation in Chinese hamster ovarian cells. Biotechnol Prog 26(5):1417–1423.

20. Van Buren N, Rehder D, Gadgil H, Matsumura M, Jacob J. 2008. Elucidation of two major aggregation pathways in an IgG2 antibody. J Pharm Sci 98(9):3013–3030.

21. Shukla AA, Gupta P, Han X. 2007. Protein aggregation kinetics during protein A chromatography. Case study for an Fc-fusion protein. J Chromatogr A 1171(1–2):22–28.

22. Fast JL, Cordes AA, Carpenter JF, Randolph TW. 2009. Physical instability of a therapeutic Fc fusion protein: Domain contributions to conformational and colloidal stability. Biochemistry 48:11724–11736.

23. Xu X, Didio DM, Leister KJ, Ghose S. 2010. Disaggregation of high-molecular weight species during downstream processing to recover functional monomer. Biotechnol Prog 26(3):717–726.

24. Van Maarschalkerweerd A, Wolbink GJ, Stapel SO, Jiskoot W, Hawe A. 2011. Comparison of analytical methods to detect instability of etanercept during thermal stress testing. Eur J Pharm Biopharm 78:213–221.

25. Mitchell D, Pobre EG, Mulivor AW, Grinberg AV, Castonguay R, Monnell TE, Solban N, Ucran JA, Pearsall RS, Underwood KW, Seehra J, Kumar R. 2010. ALK1-Fc inhibits multiple mediators of angiogenesis and suppresses tumor growth. Mol Cancer Ther 9(2):379–388.

26. Li Y, Weiss WF, Roberts CJ. 2009. Characterization of high-molecular-weight nonnative aggregates and aggregation kinetics by size-exclusion chromatography with inline multi-angle laser light scattering. J Pharm Sci 98(11):3997–4016.

27. Cooper CA, Gasteiger E, Packer N. 2001. GlycoMod—A software tool for determining glycosylation compositions from mass spectrometric data. Proteomics 1:340–349.

28. Kornfeld R, Kornfeld S. 1985. Assembly of aspargine-linked oligosaccharaides. Ann Rev Biochem 54:631–664.

29. Hawe A, Kasper JC, Friess W, Jiskoot W. 2009. Structural properties of monoclonal antibody aggregates induced by freeze-thawing and thermal stress. Eur J Phram Sci 38(2):79–87.

30. Pawel-Rammingen U, Johansson BP, Bjorck L 2002. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. EMBO J 21(7):1607–1615.

31. Wang C, Eufemi M, Turano C, Giartosio A. 1996. Influence of the carbohydrate moiety on the stability of glycoproteins. Biochemistry 35(23):7299–7307.

32. Mer G, Hietter H, Lefevre JF. 1996. Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated proteinase inhibitor. Nat Struct Biol 3(1):45–53.

33. Gervais V, Zerial A, Oschkinat H. 1997. NMR investigations of the role of the sugar moiety in glycosylated recombinant human granulocyte-colony-stimulating factor. Eur J Biochem 247:386–395.

34. Roberts CJ. 2007. Non-native protein aggregation kinetics. Biotechnol Bioeng 98(5):927–938.

35. Morris AM, Watzky MA, Finke RG. 2009. Protein aggregation kinetics, mechanism, and curve-fitting–A review of the literature. Biochim Biophys Acta 1794(3):375–397.

36. Frand AR, Cuozzo JW, Kaiser CA. 2000. Pathways for protein disulfide bond formation. Trends Cell Biol 10(5):203–210.

37. Zhang W, Czupryn MJ. 2002. Free sulfhydryl in recombinant monoclonal antibodies. Biotechnol Prog 18(3):509–513.

38. Remmele RL, Callahan WJ, Krishnan S, Zhou L, Bondarenko PV, Nichols AC, Kleemann GR, Pipes GD, Park S, Fodor, Kras E, Brems DN. 2006. Active dimer of epratuzumab provides insight into the complex nature of an antibody aggregate. J Pharm Sci 95(1):126–145.