Aggregation Analysis of Therapeutic Proteins, Part 1

General Aspects and Techniques for Assessment

Tsutomu Arakawa, John S. Philo, Daisuke Ejima, Kouhei Tsumoto, and Fumio Arisaka

Many proteins are marginally stable in solution, undergoing conformational changes due to various stresses during purification, processing and storage. Such stresses can include elevated temperature, shear strain, surface adsorption, and high protein concentration (**1**). Structurally altered proteins have a strong tendency to aggregate, often leading to eventual precipitation (**2**). Irreversible aggregation is a major problem for long-term storage stability of therapeutic proteins and for their shipping and handling.

Aggregation problems have been implicated in adverse reactions and other safety issues since the beginning of clinical applications of protein pharmaceuticals. Immunoglobulin aggregates have long been known to cause anaphylactoid reactions (**3, 4**). In recent years, a serious concern about aggregation was raised by an upsurge in incidents of pure red-cell aplasia (PRCA) in patients receiving recombinant erythropoietin, although it appears that aggregation is not involved in PRCA (**5, 6**). To minimize such risks from therapeutic proteins in clinical applications, formulations must be optimized to reduce aggregation during storage, handling, and shipping (**7**).

Size-exclusion chromatography (SEC) has been a workhorse for detecting and quantifying protein aggregation (**8, 9**). But SEC is often questioned because of possible loss of proteins (soluble aggregates, in particular) by their nonspecific binding to the columns (**10**). Native gels have also been used to observe protein aggregation, but only qualitatively. Column-free techniques such as analytical ultracentrifugation (AUC), field-flow fractionation (FFF), and dynamic light scattering (DLS) now find increasing application in aggregation analysis. Here, we review advantages and disadvantages of various techniques for assessing protein aggregation. In part 1, we discuss general aspects of protein aggregation and present two techniques mentioned above: native gel electrophoresis and SEC.

Product Focus: Proteins

Process Focus: Analysis and downstream processing

Who Should Read: QA/QC, process development, manufacturing, and analytical personnel

Keywords: Electrophoresis, size-exclusion chromatography, aggregation

Level: Intermediate

Mechanisms of Aggregation

Mechanisms of protein aggregation are still not fully understood, despite the fact that aggregation is a major problem in therapeutic-protein development. One plausible mechanism is that aggregation is driven or catalyzed by the presence of a small amount of a contaminant.

ANDREI TCHERNOV (WWW.ISTOCKPHOTO.COM)

That contaminant could be a damaged form of the protein product itself, host cell proteins, or even nonprotein materials such as leachates or silica particles (Figure 1A, pathway 1).

If the contaminant is the damaged protein, then its aggregation may lead to soluble oligomers, which become larger aggregates, visible particulates, or insoluble precipitates. Such soluble oligomers, host-cell contaminants, or nonprotein materials may serve as a nucleus onto which native proteins assemble and are incorporated into larger aggregates (Figure 1A, pathway 2) (**11**). Damaged forms of a protein product can arise from chemical modification (such as oxidation or deamidation) and from conformationally damaged forms arising from thermal stress, shear, or surface-induced denaturation. Minimizing protein aggregation from this mechanism requires both chemical and physical homogeneity; that is, chemically modified or conformationally altered proteins must be removed from the final product.

A second mechanism begins with partial unfolding of the native protein

during storage (Figure 1A, pathway 2). Protein conformation is not rigid — the structure fluctuates around the time-averaged native structure to different extents depending on environmental conditions. Some partially or fully unfolded protein molecules are always present at equilibrium in all protein solutions, but most such molecules simply refold to their native structure. However, those unfolded proteins may instead aggregate with other such molecules or may be incorporated into an existing aggregate nucleus, eventually to form larger aggregates as described above. Factors such as elevated temperature, shaking (shear and air-liquid interface stress), surface adsorption, and other physical or chemical stresses may facilitate partial unfolding.

A third aggregation mechanism is reversible self-association of the native protein to form oligomers. The content of such reversible aggregates will change with total protein concentration, according to the law of mass action. The tendency of different proteins to reversibly associate is highly variable, and the strength of that association typically varies significantly with solvent conditions such as pH and ionic strength. In principle, these reversible oligomers will dissociate completely as the therapeutic protein gets highly diluted after delivery in vivo; consequently, they are somewhat less of a concern than irreversible aggregates. However, such reversible oligomers often eventually become irreversible (they are a first step along a pathway to irreversible aggregation). Thus, preventing accumulation of irreversible aggregates may require minimizing the reversible association. Further, reversible self-association can significantly alter important pharmaceutical properties of product solutions such as solution viscosity (**12**).

Detection of reversible aggregates can be especially challenging. One reason is simply that such aggregates can dissociate from dilution during the measurement process. A second, less obvious reason is that the results of any analysis method involving separation may depend on the kinetic rates of the reversible association–dissociation reactions as well as the equilibrium constants.

One consequence of the complexities of aggregate formation processes is the difficulty of linking the effect (presence of aggregates) to its underlying cause, particularly because the key damage may occur at a time or place quite separated from the observed consequence. One example arises during large-scale production of therapeutic monoclonal antibodies (MAbs), a key step in successful product development. Acid stability plays a major role in aggregation of MAbs because their purification process typically involves both low-pH elution from protein-A affinity columns and acid-treatment for viral inactivation (Figure 1B) (**13, 14**).

Low-pH exposure of MAbs results in small but significant conformational changes that depend on the pH, temperature, and solvent composition. Although such partially unfolded MAbs may not aggregate at low pH, they often do aggregate during subsequent manufacturing steps involving changes in pH or ionic strength (**15**). A larger conformational change at low pH generally leads to more aggregates upon pH increase. Typically, aggregate formation from the low-pH structure is not a fast process, but it occurs slowly from association of damaged monomers (as depicted in Figure 1B) that have not returned to their fully native structure. Indeed, this and other types of aggregation problems may not become apparent until months after manufacturing a particular lot of protein or until later stages of product development. Regardless of the mechanism of aggregation, preventing aggregation problems requires sensitive and reliable technologies for quantitative determination of aggregate content and aggregate characteristics.

Aggregation Analysis

Aggregation analysis of proteins can be formally classified into three experimental modes. As shown in Figure 2 (top panel), the first mode is the most conventional approach: A small volume of sample is applied to a separation medium and forms a band or zone. As the band migrates through the medium, the proteins separate according to differences in size, electrophoretic charge, or mass. Gel electrophoresis, SEC, FFF, and the seldom-used band sedimentation belong to this class. The movement of the band or zone is monitored using absorbance or refractive index detection.

In the second mode of analysis, initially the sample uniformly fills a measurement cell (middle panel). When an electrical or centrifugal

Figure 1: Pathway of protein aggregation

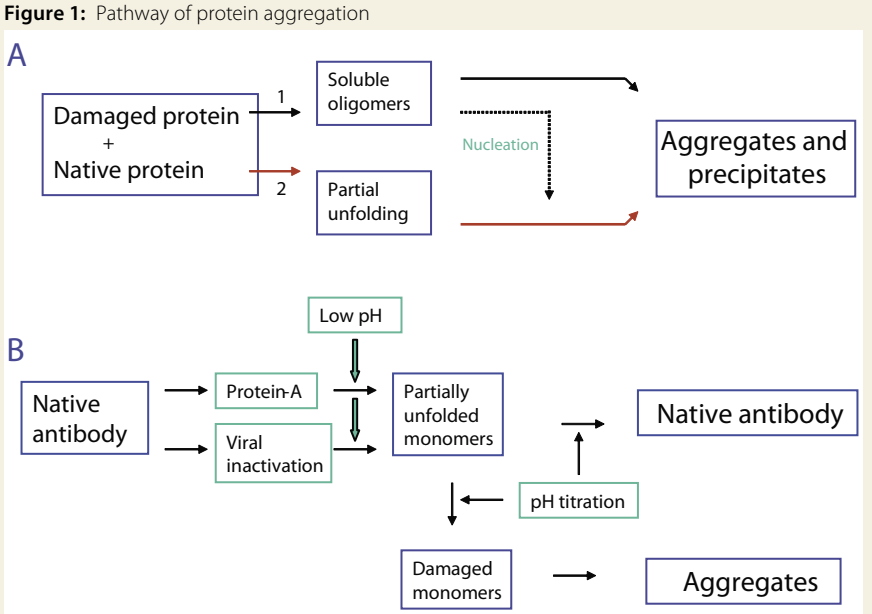


Figure 2: Classification of various techniques to analyze protein aggregation

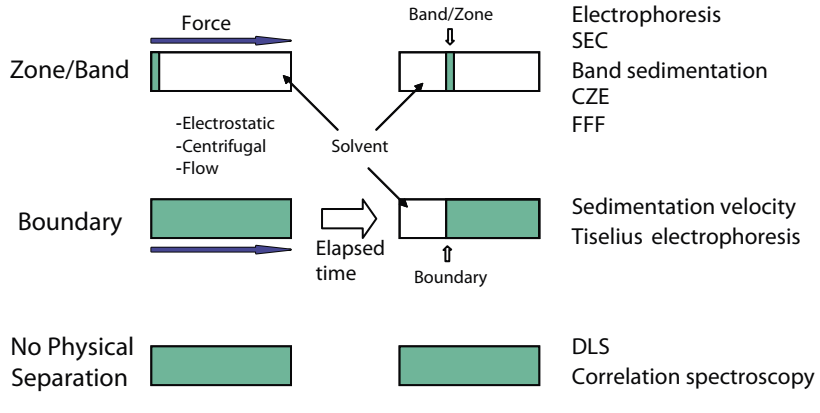
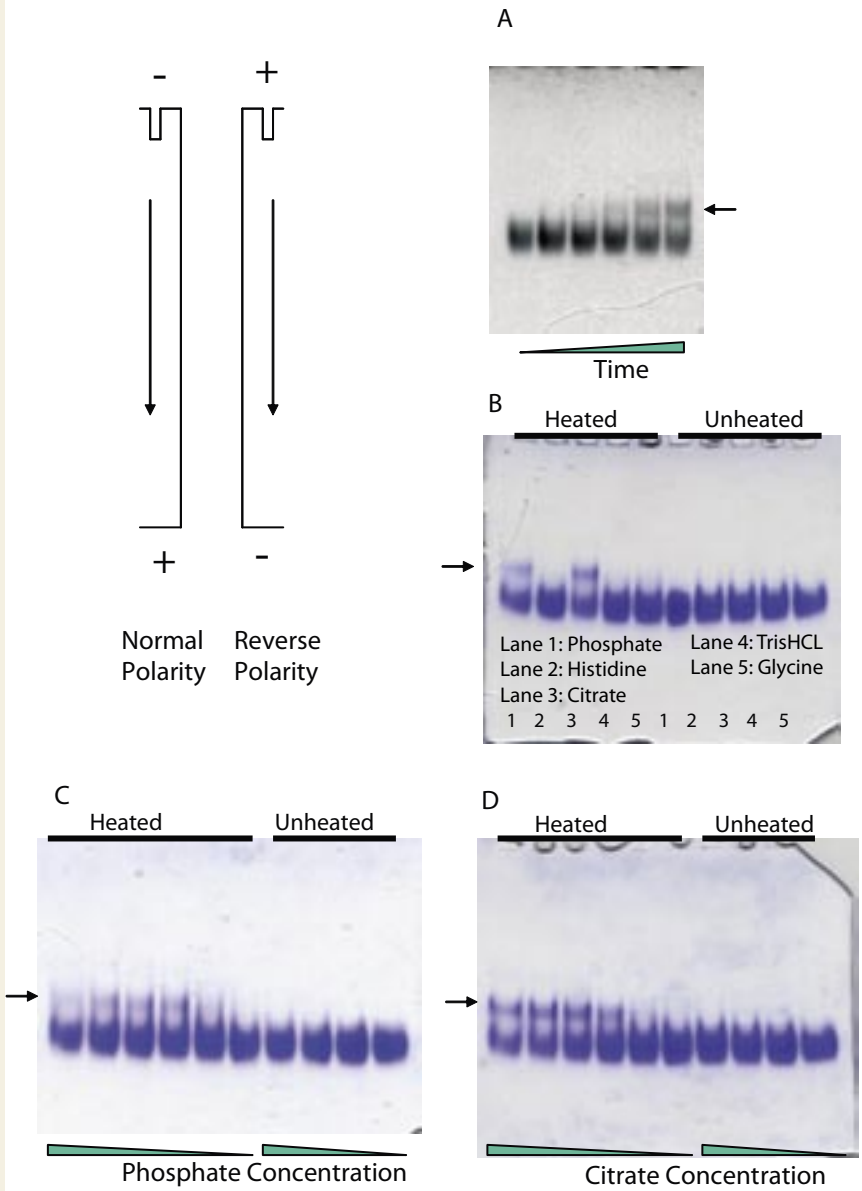


Figure 3: Native gel electrophoresis of EPO — A, heated at 79 °C with time in 20 mM phosphate; B, heated at 79 °C for 70 min in different buffers at 20 mM; C, heated at 79 °C for 70 min as a function of phosphate concentration; D, heated at 79 °C for 70 min as a function of citrate concentration



driving force is then applied, the protein moves along the applied field, leaving protein-depleted solvent, which creates a boundary between protein-free and protein-containing solution phases. The movement of this boundary over time is measured. This mode of separation is used in sedimentation-velocity and moving-boundary electrophoresis (an electrophoresis originally developed by Tiselius) (**16**).

The third category is a measurement of particle size with no physical separation (bottom panel). It is termed *correlation spectroscopy* and measures fluctuation of particles in solution due to Brownian motion (diffusion coefficients). Fluctuations of scattered light and of fluorescence intensity have been used. The most widely used method in this category is dynamic light scattering.

Native Gel Electrophoresis: Native gel electrophoresis is a convenient way to observe aggregation. For native gels, sodium dodecyl sulfate (SDS) is not used. The proteins retain their native conformation, and their mobility is governed by the ratio of electric charge to hydrodynamic friction. Hence, native gels neither denature nor dissociate proteins. Acidic proteins, which have isoelectric points below 8, can be run in the regular Tris-glycine Laemmli system. For basic proteins, native gels are more problematic and require optimization of conditions, including the type of buffer and gel, the pH, and the length of electrophoretic run. Such gels are run with the polarity reversed from that of standard SDS gels and native gels for acidic proteins. Because basic proteins are positively charged, they migrate toward the cathode as shown in Figure 3 (far left panel). For acidic proteins or on SDS-PAGE, proteins are negatively charged and migrate toward the anode.

Electrophoretic mobility of proteins in the native gel depends on both the electric charge and the hydrodynamic size. It therefore gives information on aggregation, provided that the charged state of the oligomers or aggregates is identical to that of the monomer. Figure 3 shows native gel analysis of

erythropoietin (EPO) that was expressed in mammalian cells (CHO cells) and hence is heavily glycosylated. Aggregation of EPO was induced by high-temperature treatment (heat stress). The sample in Figure 3A was heated at 79 °C in 20 mM phosphate for various times and brought to room temperature for the native gel analysis. Only a single band is observed in the unheated control (lane 1), suggesting that the EPO preparation contains no aggregates. However, a native gel with Coomassie blue staining would not detect 1–2% of small oligomers, reflecting the sensitivity limit of such staining. No change was seen in the mobility after heating at 79 °C for 5–15 minutes (lanes 2 and 3), indicating no apparent formation of aggregates within the limit of Coomassie blue staining.

Because EPO is completely unfolded at 79 °C (well above the melting temperature of EPO) within a few minutes, the observed full recovery of the native protein indicates that EPO readily regains the native structure upon cooling (**17, 18**). After 30 minutes of heating (lane 4), a small amount of low mobility band corresponding to dimers is observed, as indicated by the arrow (Figure 3A). As the incubation time is increased, more dimers and oligomers are formed and the monomer decreases (lanes 5 and 6). The observed formation of dimers and oligomers indicates that a prolonged incubation causes unfolding and consequent aggregation of EPO, which cannot be reversed by low temperature, an aggregation pathway described in Figure 1A (pathway 2). The formation of dimers, some of which are covalently linked by disulfide bonds, and higher oligomers has been observed upon incubation at elevated temperatures (data not shown; see also **19**). There is no dimer formation when heated at 65 °C up to 90 minutes (data not shown), which is still above the melting temperature of EPO, indicating again that EPO readily refolds to monomer.

We tested other buffers (citrate, TrisHCl, histidine, and glycine) all at 20 mM for 70 minutes for their effects on the reversibility of thermal unfolding at 79 °C. The results in Figure 3B show that these buffers have no effect on the native protein (before heating), whereas TrisHCl, histidine, and glycine buffers gave a nearly complete protection from aggregation. Sedimentation velocity gave more quantitative and detailed information on the heat-induced aggregation of EPO as will be described later (Part 2).

Next, native gel analysis was applied to EPO samples heat-stressed in different concentrations of phosphate (Figure 3c). It is evident that there are more soluble oligomers at intermediate phosphate concentrations. However, because native gels are only semiquantitative it is not clear whether monomer content is actually increasing at higher phosphate concentration, e.g., 0.2 M (last lane), which illustrate a major drawback of this approach. Figure 3D shows the effect of citrate concentration. The amounts of soluble oligomers and monomers reach a plateau at intermediate citrate concentrations, above which there appears to be no further increase in aggregates. These native gel results indicate that EPO aggregation occurs after thermal unfolding and depends on solvent conditions. Although this approach is not quantitative and does not fully reveal the nature of the aggregates, it is a useful screening tool to find solvent conditions that may reduce aggregation.

Although natural and CHO-derived EPO are highly soluble in aqueous buffer because of heavy glycosylation, unglycosylated EPO is prone to aggregation because of its high content of nonpolar amino acids (**17, 19**). Heat-induced aggregation of EPO as observed above may therefore be due to exposure of hydrophobic regions of the protein.

Native gels of purified monoclonal antibodies (MAb-a and MAb-b) show a cluster of sharp bands due to heterogeneous glycosylation, which results in either charge or size heterogeneity or both (Figure 4); each band appears to correspond to a homogeneous protein species. These purified MAb samples contain about 3% aggregates as determined by sedimentation velocity. Within the detection limit of silver staining, no bands are observed above these major components, indicating little aggregation, consistent with the sedimentation velocity analysis.

Native gel analysis was also used to analyze aggregates of these MAbs. Aggregates were generated by subjecting the purified MAbs to low pH and high temperature (stressed samples). The heat-stressed MAbs thus obtained showed several bands above the major bands (highlighted by a blue bracket), indicating that heat-treatment generated aggregates. Those bands corresponding to aggregated species are more diffuse than the monomer band, suggesting that the aggregate bands are heterogeneous. Although native gels can separate multiple forms of monomer antibodies through heterogeneous glycosylation, this technique is not as quantitative as other techniques described here.

Size-Exclusion Chromatography (SEC): As mentioned above, SEC is the workhorse for aggregation analysis of pharmaceutical proteins. Aggregation analysis of pharmaceutical proteins using this technique is an absolute nearly always required for regulatory approval. Figure 5 shows the resolution of analytical SEC for MAb-a using a G3000SWXL HPLC column (Tosoh Biosciences) and an elution solvent of 0.1 M phosphate, pH 6.8, in the absence and presence of 0.2 M arginine. To generate aggregates, the MAb sample was subjected to low pH and heat treatment (in the absence of arginine) as described above. Panel A corresponds with the SEC results of the unstressed samples, whereas panel B shows the results of the heat-stressed samples. The nonstressed MAb-a shows a sharp peak corresponding with the monomeric form of the sample. Only about 2% total aggregates are present in the sample.

It is well recognized, however, that the SEC column matrix (stationary phase) can nonspecifically bind proteins, and this is especially true for aggregates and, hence, makes aggregation analysis ambiguous (**10**). We have shown that adding arginine greatly reduces such

nonspecific binding (**20**). Figure 5A (right panel) shows the effects of adding 0.2 M arginine to the elution solvent. Although the total aggregate content is similar (~2%), there is a clear increase in one of the aggregate species (indicated by arrow), suggesting that such aggregates, although present in small amount, are lost during SEC when arginine is not present; there is no decrease in the monomer recovery in the presence of 0.2 M arginine. Figure 5B shows the SEC results for heat-stressed MAb-a. It is evident that the aggregate peaks are much larger than those seen for the nonstressed MAb-a (panel A), with the aggregate content increasing to ~26% (from 2%). However, the heat-stressed sample actually contains many more aggregate than the above value, as also shown in Figure 5B (right panel). In the presence of 0.2 M arginine, the aggregate peaks are much larger, with the aggregate content at 43%, indicating again that a large number of aggregates are lost during the SEC unless arginine is present. It should be noted that the observed increase in aggregate content is not due to formation of aggregates by arginine during SEC. First, the amount of monomer eluting is not decreased by the arginine. Second, in our limited experience, arginine does not denature, nor does it enhance aggregation for several proteins tested (**21**).

Obviously, careful monitoring of total sample recovery during chromatography should reveal the poor recovery of aggregates such as that in Figure 5B and serve as a warning that the SEC method is not working properly. However, even very careful monitoring of total recovery may not reveal subtle differences such as those shown in Figure 4A.

Figure 4: Native gel electrophoresis of MAb, 7.5 %T homogeneous gel (Phastgel); purified, 1 µg; heated, 2 µg; silver staining

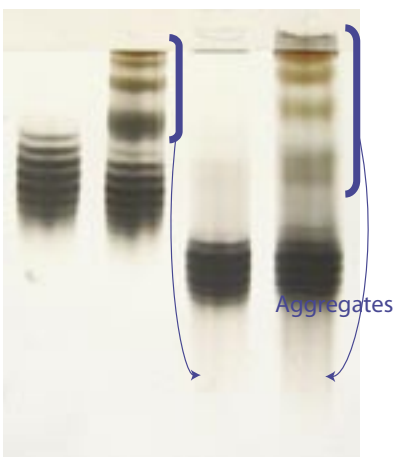


Figure 6 shows the SEC results for MAb-b. The unstressed sample contains only ~1% aggregates whether the SEC is done in the absence or presence of 0.2 M arginine (panel A). Peaks corresponding to aggregates are much larger for the stressed sample (panel B): ~38% in the absence of arginine and ~43% in the presence of 0.2 M arginine, indicating again that arginine suppresses nonspecific binding of aggregated antibodies, leading to a higher recovery of the aggregates. It appears that the shoulder of the aggregate peak (indicated by arrow) is higher in the presence of arginine, suggesting that larger aggregates, which elute earlier, tend to stick to SEC columns in the absence of arginine. Of course, it is uncertain whether arginine completely suppresses the loss of aggregates on the column.

It is thus evident that inclusion of arginine in the SEC mobile phase suppresses protein adsorption to the silica surface. This is most likely due to preferential binding of arginine to the unbound protein, which is in equilibrium with the surface bound form (**22, 23**). As in Figure 7, arginine shifts the equilibrium binding of antibodies toward the dissociated state. Can arginine then dissociate the bound proteins from the column? The answer is no, at least for the MAb-a tested. As depicted in Figure 7, the protein molecules bound to the SEC column surface often undergo conformational changes, leading to binding-elution hysteresis or irreversible binding (**24**). Arginine is incapable of dissociating these stronger interactions.

Another important advantage of arginine is that it reduces lot-to-lot manufacturing variation of SEC columns in aggregation analysis, as illustrated in Figure 8. Three lots of Superdex-75 (GE Healthcare, www.amershambiosciences.com) show a varying degree of aggregate content for interleukin-6. An identical antibody sample was applied to the three different lots of a same column type. As shown in panel A, little aggregate is eluted from this column, although the sample contains a significant number of aggregates. The column lot shown in panel C showed a much larger aggregate peak than the other two lots (panel A and B), indicating that nonspecific adsorption of proteins varies with different lots. When 0.25 M arginine was included, this variability disappeared. As shown in three panels, with arginine the aggregate peaks are all sharp and similar in size, indicating that arginine suppresses nonspecific protein adsorption, independent of the lots. The ratio of the total aggregate content measured in the absence or presence of 0.25 M arginine is ~0 for lot 0314065 (no distinct peak observed in the absence of arginine), 0.37 for lot 0437127, and 0.95 for lot 0511099. That is, the last lot is least sticky, at least for these aggregates. Thus, inclusion of arginine in the elution solvent eliminates the need to reoptimize elution conditions when changing to a new lot of columns.

The molecular weight of aggregates can be determined using an on-line light-scattering detector, which measures the amount of light scattered by the proteins. When combined with the protein concentration data from UV absorbance or refractive index detectors, the light scattering can give the average molecular weight for each peak eluting from an SEC column. The molecular weight of the aggregates relative to the monomer indicates the stoichiometry of the aggregates formed.

Looking Ahead

As shown here, the mechanisms of protein aggregation can introduce problems. Several technologies can be used to determine that level of aggregation, among which are native gel electrophoresis and SEC. In part 2, we will introduce three other technologies: sedimentation velocity,

Figure 5: SEC analysis of MAb-a

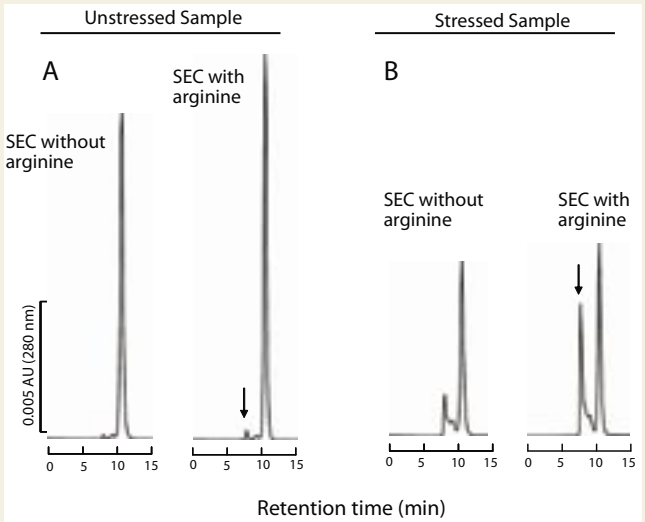
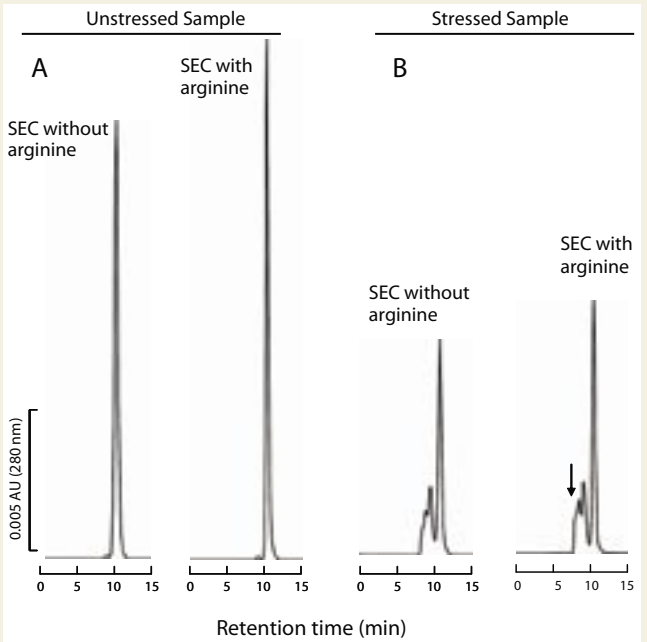


Figure 6: SEC analysis of MAb-b



FFF, and DLS, which are all free from the column matrix.

References

1 Manning MC, Patel K, Borchardt RT. Stability of Protein Pharmaceuticals. *Pharm. Res.* 6(11) 1989: 903–917.

2 Chen B-L, et al. Strategies to Suppress Aggregation of Recombinant Keratinocyte Growth Factor During Liquid Formulation Development. *J. Pham. Sci.* 83(12) 1994: 1657–1661.

3 Ring J, Stephan W, Brendel W. Anaphylactoid Reactions to Infusions of Plasma Protein and Human Serum Albumin: Role of Aggregated Proteins and of Stabilizera Added During Production. *Clin. Allergy* 9(1) 1979: 89–97.

4 Ring J, et al. Anaphylactoid Reactions Due to Non-Immune Complex Serum Protein Aggregates. *Monogr. Allergy* 12, 1977: 27–35.

5 Frost H. Antibody-Mediated Side Effects of Recombinant Proteins. *Toxicology* 209(2) 2005; 155–160.

6 Smalling R, et al. Drug-Induced and Antibody-Mediated Pure Red Cell Aplasia: A Review of Literature and Current Knowledge. *Biotechnol. Annu. Rev.* 10, 2004: 237–250.

7 Arakawa T, Kita Y, Carpenter JF. Protein-Solvent Interactions in Pharmaceutical Formulation. *Pharm. Res.* 8(3) 1991: 285–291.

8 Wang W. Instability, Stabilization, and Formulation of Liquid Protein Pharmaceuticals. *Int. J. Pharm.* 185(2) 1999: 129–188.

9 Jones AJS. Analysis of Polypeptides and Proteins. *Adv. Drug Deliv. Rev.* 10, 1993: 29–90.

10 Stulik K, Pacakova V, Tich, M. Some Potentialities and Drawbacks of Contemporary Size Exclusion Chromatography. *J. Biochem. Biophys. Methods* 56(3) 2003: 1–13.

11 Chi EY, et al. Heterogeneous Nucleation-Controlled Particulate Formation of Recombinant Human Platelet-Activating Factor Acetylhydrolase in Pharmaceutical Formulation. *J. Pham. Sci.* 94(2) 2005: 256–274.

12 Shire SJ, Shahrokh Z, Liu J. Challenges in the Development of High Protein Concentration Formulations. *J. Pharm. Sci.* 93(6) 2004: 1390–1402.

13 Vermeer AWP, Norde W. The Thermal Stability of Immunoglobulin: Unfolding and Aggregation of a Multi-Domain Protein. *Biophys. J.* 78(1) 2000: 394–404.

14 Calmettes P, Cser L, Rajnavolgyi E. Temperature and pH Dependence of Immunoglobulin G Conformation. *Arch. Biochem. Biophys.* 291(2) 1991: 277–283.

15 Ejima D, et al. Effects of Acid Exposure on the Conformation, Stability, and Aggregation of Monoclonal Antibodies. *Proteins: Structure, Function and Bioinformatics.* In press.

16 Tiselius A. Electrophoresis of Serum Globulin. *I. Biochem. J.* 31(2), 1937: 313–317.

17 Narhi LO, et al. The Effect of Carbohydrate on the Structure and Stability of Erythropoietin. *J. Biol. Chem.* 266(34) 1991: 23022–23026.

18 Arakawa T, Philo JS, Kita Y. Kinetic and Thermodynamic Analysis of Thermal Unfolding of Recombinant Erythropoietin. *Biosci. Biotechnol. Biochem.* 65(6) 2001: 1321–1327.

19 Endo Y, et al. Heat-Induced Aggregation of Recombinant Erythropoietin in the Intact and Deglycosylated States As Monitored By Gel Permeation Chromatography Combined with a Low-Angle Laser Light Scattering. *J. Biochem.* 112(5) 1992: 700–706.

20 Ejima D, et al. Arginine As an Effective Additive in Gel Permeation Chromatography. *J. Chromatogr. A* 1094(1-2) 2005: 49–55.

21 Tsumoto K, et al. Role of Arginine in Protein Refolding, Solubilization, and Purification. *Biotechnol. Prog.* 20(5) 2004: 1301–1308.

22 Kita Y, et al. Contribution of the Surface Free Energy Perturbation to Protein-Solvent Interactions. *Biochemistry* 33(50) 1994: 15178–15189.

23 Lin TY, Timasheff SN. On the Role of Surface Tension in the Stabilization of Globular Proteins. *Protein Sci.* 5(2) 1996: 372–381.

24 Jennissen HP, Botzet G. Protein Binding to Two-Dimensional Hydrophobic Binding-Site Lattices: Adsorption Hystresis on Immobilized Butylresidues. *Int. J. Biolog. Macromolecules* 1, 1979: 171–179.

Corresponding author ***Tsutomu Arakawa*** *is president of Alliance Protein*

Laboratories, Thousand Oaks, CA 91360, tarakawa2@aol.com; ***John S. Philo*** *is vice president and director of biophysical chemistry at Alliance Protein Laboratories;* ***Daisuke Ejima*** *is a chief chemist at AminoScience Laboratories, Ajinomoto Co., Inc., Kawasaki, Kanagawa 210-8681, Japan;* ***Kouhei Tsumoto*** *is an associate professor at the Department of Medical Genome Sciences, The University of Toyko, Kashiwa 277-8562, Japan; and* ***Fumio Arisaka*** *is an associate professor in the Department of Molecular Bioprocessing, Tokyo Institute of Technology, Yokohama 226-8501, Japan.*

Figure 7: Non-specific protein adsorption to surface

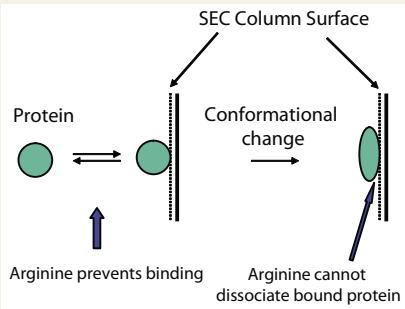


Figure 8: SEC analysis of aggregation of interleukin-6; an interleukin-6 monomer–dimer mixture (13 µg) sample with 0.1 M Na phosphate buffer, pH 6.8, containing 0.25 M NaCl or arginine