Published in final edited form as:
J Pharm Sci. 2018 May ; 107(5): 1313–1321. doi:10.1016/j.xphs.2018.01.015.

Flow microscopy imaging is sensitive to characteristics of subvisible particles in peginesatide formulations associated with severe adverse reactions

Austin L. Daniels and **Theodore W. Randolph**
Center for Pharmaceutical Biotechnology, Dept. of Chemical and Biological Engineering, University of Colorado, Boulder, Colorado 80309-0596

Abstract

The presence of subvisible particles in formulations of therapeutic proteins is a risk factor for adverse immune responses. Although the immunogenic potential of particulate contaminants likely depends on particle structural characteristics (e.g., composition, size, and shape), exact structure-immunogenicity relationships are unknown. Images recorded using flow imaging microscopy reflect information about particle morphology, but flow microscopy is typically used to determine only particle size distributions, neglecting information on particle morphological features that may be immunologically relevant. We recently developed computational techniques that utilize the Kullback-Leibler divergence and multidimensional scaling to compare the morphological properties of particles in sets of flow microscopy images. In the current work, we combined these techniques with expectation maximization cluster analyses, and used them to compare flow imaging microscopy datasets that had been collected by the US FDA after severe adverse drug reactions (including seven fatalities) were observed in patients that had been administered some lots of peginesatide formulations. Flow microscopy images of particle populations found in the peginesatide lots associated with severe adverse reactions in patients were readily distinguishable from images of particles in lots where severe adverse reactions did not occur.

Keywords

Protein Formulation; Protein Aggregation; Image Analysis; Regulatory Science

Introduction

Protein therapeutics offer numerous clinical benefits, and now comprise the fastest-growing class of drugs.[1] A challenge in the development of protein therapeutics is that they may elicit adverse drug reactions (ADRs) which include acute responses such as anaphylaxis during IV administration, or long-term adverse reactions such as immune responses wherein patients

---

produce anti-drug antibodies (ADAs)[2–4]. The majority of current protein therapeutics are immunogenic in at least some patients[5], and in some cases (e.g., interferon beta[6–8]) adverse immune responses may be observed in up to half of patients treated, reducing efficacy.[8] Adverse immune responses can result in clinical trial failures.[9]

There are a number of potential causes and risk factors associated with ADRs against protein therapeutics.[10–17] Among these risk factors is the presence of aggregates within protein formulations.[11, 18–28] Numerous animal studies[29], human clinical studies[30–34] and *in vitro* studies[35–37] have associated particulate contaminants with infusion reactions, anaphylaxis, and activation of the innate and adaptive immune system.[38, 39] Aggregation occurs as a result of various stresses to which proteins may be exposed, and different stresses such as freeze-thawing, exposure to air-water interfaces, pH extremes, elevated temperatures or chemical degradation produce different distributions of aggregates that are polydisperse in size and morphology.[40] *In vivo*, these aggregate populations may provoke different levels and types of immune responses.[40,20,36] For example, in one study, protein aggregates produced by process-related conditions and low pH were not immunogenic.[41] Another study showed that larger, insoluble aggregates found in an antibody formulation after UV-light exposure were more immunogenic than soluble oligomeric aggregates of the same protein.[42] At the present time, it is unclear which characteristic(s) of protein aggregates dictate their immunogenicity, in part because of the difficulties involved analyzing the different populations of particles (e.g., particles generated through different mechanisms of formation) that may be present in a given sample. Better techniques for characterizing aggregates are necessary in order to identify the features of protein aggregates that influence their ability to trigger ADRs upon administration—features that could be then monitored to assess the risk of ADRs and allow the most dangerous aggregate populations to be identified and prioritized for removal.

Protein drug manufacturers frequently use flow-imaging microscopy (FIM) to monitor the concentrations of micron-sized subvisible particles (e.g., protein aggregates, silicone oil droplets, air bubbles) present in protein formulations. In this technique, a sample is pumped through a microfluidic channel where a microscope records digital images of particles of size greater than about 2 μm. This technique yields either grayscale or color images representative of the 10^4– 10^6 individual particles larger than 2 μm typically present in a given sample. These image datasets are frequently large, with data file sizes of up to a gigabyte per sample. These collections of images potentially offer a wealth of particle structural information, but FIM is frequently used to obtain only particle size distributions as a histogram for a given sample. While convenient, this practice neglects other potentially relevant morphological features that could be extracted from these images. We hypothesize that the neglected information about particle morphology contained in the rich data sets generated by FIM could be relevant to determining risk of ADRs from particles within a protein formulation.

We recently developed a technique[43] to analyze collections of FIM datasets in order to differentiate between various populations of particles represented in the datasets. In this technique, the distributions of particle properties in each sample are compared to the distributions of properties in other samples via the symmetrized Kullback-Leibler divergence

(KLD). A matrix of pairwise values of this divergence can then be processed via multidimensional scaling (MDS) to obtain a low-dimensional embedding of the data that captures the relative similarity between one dataset and the others included in the analysis. We previously demonstrated[43] that this technique can successfully differentiate between populations of particles formed in monoclonal antibody solutions that had been subjected to different aggregation-inducing stresses (freeze-thawing, shaking and pH changes, and elevated temperatures).

A recent study by the US FDA associated elevated levels of nano- and microparticles found in a marketed formulation of peginesatide (Omontys®; Affymax, Inc., Cupertino, CA) with severe ADRs in patients.[44] The drug, an erythropoiesis-stimulating agent consisting of a covalently dimerized synthetic peptide linked to polyethylene glycol, received FDA approval in 2012 for two formulations: a single-use vial formulation (SUV) and multi-use vial (MUV) formulation. The two formulations contain peginesatide at the same concentration, but have different excipients.[45] Although the SUV formulation was used predominantly during the clinical trials, only the MUV formulation was marketed. The marketed MUV formulation was linked to 49 cases of anaphylaxis (7 of which were fatal) and a hypersensitivity rate of 3.5 per 1000 exposed patients—significantly higher than the 0.84 per 1000 exposed patients rate that had been noted for the SUV formulation during the clinical trials. After the product was recalled voluntarily, the FDA investigated both the SUV and MUV formulations, conducting a variety of analyses in search of potential causes of the severe ADRs. Standard testing of the SUV and MUV formulations revealed that both formulations conformed to product specifications, including the pharmacopeial limitations on the concentrations and size distributions for particles described by USP <788>. However, although both formulations met current limitations on particle content, it was discovered that the marketed MUV formulation had higher and more variable concentrations of subvisible particles than did the SUV formulation. Direct causality could not be established, but the analysis conducted by the FDA found that elevated subvisible particle content in the MUV formulations compared to that in the SUV formulations was associated with the observed increased hypersensitivity reactions seen for the marketed MUV formulation of peginesitide.[44]

The FDA study[44] found that higher particle levels in MUV formulations of Omontys® were associated with increased rates of ADRs, but did not examine whether the formulations differed in particle characteristics other than concentration. Such differences might reflect different mechanisms by which the particles formed in the SUV and MUV formulations, which in turn could affect the propensity of the particles to generate ADRs. In the present study, we apply our KLD-MDS approach to the flow microscopy image datasets collected by the FDA in order to discern whether flow microscopy imaging can be used to differentiate between the ADR-associated particles found in MUV formulations of Omontys® and the particle populations that did not provoke adverse responses found in SUV formulations.

Materials and Methods

Materials

Intravenous immunoglobulin ("IVIg", GAMMAGARD LIQUID) was obtained from Baxter International (Deerfield, IL). 1× phosphate buffered saline (PBS) containing 144 mg/mL potassium phosphate monobasic, 795 mg/mL potassium phosphate dibasic, and 9000 mg/mL sodium chloride was obtained from Gibco (Waltham, MA). Hellmanex III was obtained from Hellma Analytics (Mullheim, Germany). All other salts and materials used in buffer preparation were reagent grade or higher.

Flow-Imaging Microscopy (FIM)

Flow-imaging microscopy datasets from the peginesatide investigation were provided by the FDA under a Freedom of Information Act (FOIA) request. In their investigation of the drug[44], the FDA analyzed samples from several SUV and MUV lots using a FlowCam VS1 system (Fluid Imaging Technologies, Inc., Scarborough, ME). The instrument used an 80-μm flow cell and a 10× objective. 450 μL of sample were analyzed in each measurement.

FIM datasets that we obtained from the FDA had been collected in three sets of FlowCam measurements referred to as "experiments" in the original study. We will use their nomenclature and denote these data collections as "Experiment A", "Experiment B" and "Experiment C". Due to limited sample volume, FIM settings were optimized over the course of data collection and thus each experiment used slightly different FIM settings. Both SUV and MUV samples were measured in each experiment. Experiment A contains 12 FIM datasets taken from 4 MUV lots and 4 datasets taken from a single SUV lot. Experiment B contains 11 datasets taken from 4 MUV lots and 8 datasets taken from 3 SUV lots. Experiment C contains 12 datasets taken from 4 MUV lots and 12 datasets taken from 3 SUV lots. Representative images taken from Experiment C for both formulations are shown in Figure 1.

Image Analysis

Data analysis was performed in Python 3.6 (Python Software Foundation, OR). Images of the particles identified via the FlowCam instrument were imported into the software and segmented using custom image processing code to identify the particle-containing regions of the image.

This analysis results in a "particle mask" or the portion of the raw image identified as a particle. The particle mask was then used to calculate several morphological properties for each particle. The area based diameter of a particle was calculated by calculating the area of the particle mask in pixels and calculating the diameter of a circle with the same area. The aspect ratio of a particle was calculated by fitting the shape of the particle mask to an ellipse using principle components analysis (PCA) and dividing the length of the minor axis by the length of the major axis. The circularity of the particle was calculated by calculating the perimeter of the particle mask and dividing by the perimeter of a circle with the same area as the particle. The average particle intensity (i.e. grayscale color) was calculated by averaging the intensity of the pixels over the area of the particle. Histograms of the various particle properties were constructed in order to visualize the distribution of properties within the datasets. The properties included in this analysis vary in scale. Aspect ratio, circularity, and average intensity values may vary from 0 to 1, but particle diameters can take on a much wider range of values. To avoid bias resulting from these differences in scale, each property was normalized by subtracting the mean value and dividing by the standard deviation before further analysis, effectively giving each observed property equal weighting in our analysis.

Kullback-Leibler Divergence (KLD)

The Kullback-Leibler Divergence is an information theory metric that can be used to estimate the similarity between two probability distribution functions (PDFs). A low value of this divergence indicates little difference in the distributions of particle properties and increasingly higher values indicate more strongly dissimilar distributions.

The KLD between two PDFs $P(\vec{d})$ and $Q(\vec{d})$ over a vector of particle properties $\vec{d}$ can be calculated by:

$$K(P \| Q) = \int P(\vec{d}) \ln\left(\frac{P(\vec{d})}{Q(\vec{d})}\right) \mathrm{d}\vec{d} \quad (1)$$

where $K(P \| Q)$ is the KLD between distributions $P(\vec{d})$ and $Q(\vec{d})$ and $\vec{d}$ is a vector containing normalized values of the four particle properties included in this analysis: diameter, aspect ratio, circularity, and average intensity. In this analysis $P(\vec{d})$ is an estimate of the distribution of particle properties in a FIM measurement containing $n_p = 2000$ particles with properties $\vec{d}_k$ where $k = 1 \ldots n_p$ indexes particles. $P(\vec{d})$ is estimated from $\vec{d}_k$ using a kernel density approximation with a hard sphere kernel of radius 0.6. $Q(\vec{d})$ is defined similarly for 2000 particles taken from a separate FIM dataset. As previously described[43], since $P(\vec{d})$ describes the frequency of specific values of $\vec{d}_k$ appearing in n_p particles, we can approximate $K(P \| Q)$ as:

$$K(P \| Q) = \frac{1}{n_p} \sum_{k=1}^{n_p} \ln\left(\frac{P(\vec{d}_k)}{Q(\vec{d}_k)}\right) \quad (2)$$

It is important to note that the KLD is not symmetric about the two distributions; $K(Q \| P)$ will yield a different value than $K(P \| Q)$ unless $P = Q$ due to the P weighting on the integral in eq. (1). This asymmetry can pose issues when trying to interpret the KLD as a measure of distribution similarity. We therefore use a symmetrized form of the KLD:

$$K(P, Q) = \frac{1}{2}(K(P \| Q) + K(Q \| P)) \quad (3)$$

where $K(P, Q)$ is the symmetrized KLD. Future mentions of the KLD will refer to this symmetrized form.

We wish to use eq. (2–3) to compare all possible pairs of FIM measurements in a dataset containing N FIM measurements. To compare dataset i to dataset j where $i = 1 \ldots N$ and $j = 1 \ldots N$ for both index datasets, we first compute P_i and Q_j, the distribution of particle properties in datasets i and j respectively. The particle properties that we utilize in this analysis are the particle diameter, aspect ratio, circularity, and average intensity, but other measures derived from the images, such as estimated masses of the individual particles[46] could also be used. We can then use these distributions to construct a matrix A whose elements a_{ij} are the squared pairwise divergences between datasets i and j. a_{ij} can be calculated as:

$$a_{ij} = K(P_i, Q_j)^2 \quad (4)$$

Multidimensional Scaling (MDS)

While matrix A contains significant information about the similarity between pairs of datasets, it is challenging to extract global similarity information from these pairwise divergence values. We can use multidimensional scaling (MDS) to find a low-dimensional embedding of the datasets that best captures the pairwise similarity information contained in A as was described and performed in previous work[43]. MDS techniques are designed to operate on a matrix of pairwise distances like the distances between cities on a map. Although the KLD is a divergence between distributions and not a distance metric, in this analysis we will treat the symmetrized KLD as a distance measurement and use MDS techniques as a method to visualize the relative similarity between datasets.

In classical multidimensional scaling (CMDS), we seek values of hypothetical points $\vec{x}_i$ where $i = 1 \ldots N$ again indexes datasets whose values are set so that the distance between $\vec{x}_i$ and $\vec{x}_j$ (where $j = 1 \ldots N$ again indexes datasets) is similar in value to the corresponding entry in matrix A a_{ij}. To perform this analysis, we first apply double centering to A using the centering matrix J whose elements j_{kl} are defined as:

$$j_{kl} = \delta_{kl} - \left(\frac{1}{N}\right) \quad (5)$$

where $k = 1 \ldots N$ and $l = 1 \ldots N$ are now indices in the centering matrix and δ_{kl} is the Kronecker delta. This matrix can be used to double center matrix A or to subtract the mean of each row and each column from the matrix:

$$B = -\frac{1}{2}JAJ \quad (6)$$

where B is the double centered matrix. The minimization to find points $\vec{x}_i$ can now be written in terms of entries in the double centered matrix b_{ij} as:

$$\min_{x} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| b_{ij} - \left\| \vec{x}_i - \vec{x}_j \right\|^2 \right\|^2 \quad (7)$$

The values $\vec{x}_i$ that minimize equation (N) can be found by performing eigendecomposition analysis on matrix B. Matrix B can be represented in terms of a matrix of N eigenvalues and N corresponding eigenvectors:

$$B = Q\Lambda Q^{-1} \quad (8)$$

where Λ is a diagonal matrix of eigenvalues and Q is a matrix of the corresponding eigenvectors. To find values two dimensional coordinates that satisfy the minimization in eq. (7), we simply select the two eigenvectors with the largest corresponding eigenvalues. $\vec{x}_i$ can then be calculated using:

$$\vec{x}_i = Q_2\Lambda_2^{1/2} \quad (9)$$

Where Λ_2 is a diagonal matrix of the two highest eigenvalues and Q_2 a matrix containing the corresponding eigenvectors. This equation yields a two-dimensional representation of the pairwise KLD values that can easily be plotted and analyzed.

KLD-MDS Plots

To compare the FIM datasets, the distribution of particle diameter, aspect ratio, circularity, and average intensity in each dataset were compared to those of the other datasets via the Kullback-Leibler divergence. The pairwise divergences were then analyzed via classical multidimensional scaling to obtain a 2D embedding for each dataset. The resulting 2D coordinates reflect the underlying similarity between datasets as measured by the KLD: two datasets that are similar as indicated by a low value of the KLD will generally appear closer together on the figure than two datasets that are more dissimilar.

It is important to note that MDS assigns coordinates to the datasets so that the distance between any two points describes the relative value of the corresponding value of the KLD. The distances between points on the 2D projection thus are the main quantities that can be used to interpret the FIM datasets described by the figure. The axes on which the coordinates are oriented are chosen to represent these distances and do not have any easily discernable physical meaning in terms of either the underlying particle properties or the values of the KLD between datasets. Since only the relative locations of the points in the embedding is significant, for ease in visualization we chose to rotate the axes obtained from KLD-MDS so that the average coordinate of both SUV and MUV datasets lay on the x-axis and that the SUV was on the left of the embedding.

KLD-MDS embeddings were obtained using data from Experiments A, B, and C individually. Additionally, we also performed the analysis using data from all three experiments at once. Since the FlowCam settings varied between the three experiments, it is possible that the differences in settings could introduce artifacts into the final KLD-MDS embedding. To account for these potential differences, data in each experiment was normalized to the mean and standard deviation of properties within that experiment before normalizing to the overall mean and standard deviation.

Clustering Analysis

KLD-MDS can be used to compare the types of particles present in the two formulations to determine whether the formulation influences the morphology of the particles present in a dataset. Since only a finite number of particles are present in a given sample, we expect some variation in the types of particles present between every pair of FIM measurements in a given dataset—even for samples taken from the same formulation. If the formulation significantly influences the morphology of the particles in the sample, two FIM measurements taken from different formulations will have more dramatic differences in particle morphology than two measurements taken from the same formulation. These differences in particle morphology will then be apparent as larger KLD values for pairs of measurements taken from different formulations than for measurements taken from the same formulation. This pattern in KLD values will manifest as clustering in the final embedding: FIM measurements on SUV samples will embed closer to other measurements performed on SUV samples and further away from MUV measurements. We can therefore identify a formulation-dependent shift in subvisible particle properties by assessing the clustering in the final KLD-MDS plots. If the samples can be successfully clustered by formulation, then the subvisible particles likely exhibit a formulation-dependent shift in particle properties indicative of a change in the subvisible particle population.

The clustering in the KLD-MDS plots can be assessed visually to determine if the datasets cluster by formulation. Alternatively, various metrics can be used to quantitatively measure the resolution of the clustering. One such metric is the silhouette coefficient[47], which is a measure of how similar a given data point in a cluster is to other points in its cluster relative to its similarity to points outside its cluster. The silhouette coefficient for a given point x_i in a set of two clusters $si(x_i)$ is calculated as

$$si(x_i) = \frac{d_{inside}(x_i) - d_{outside}(x_i)}{\max\left(d_{inside}(x_i), d_{outside}(x_i)\right)} \quad (10)$$

Where $d_{inside}(x_i)$ and $d_{outside}(x_i)$ are the average Euclidean distance between point x_i and points inside the cluster x_i belongs to and outside the cluster, respectively. The silhouette coefficient can take on values from −1 to 1 with higher values indicating increasing relative similarity to the other points in the cluster x_i is assigned to. Since every point in the KLD-MDS dataset will have a silhouette coefficient associated with it, we report the average silhouette coefficient for SUV and MUV datasets in each of the 3 experiments.

We can also identify if the known clustering in formulation can be recovered using an unsupervised clustering technique. To accomplish this, expectation maximization[48] (EM) was applied to the unlabeled coordinates from each experiment to estimate two 2D Gaussian Distributions to represent the two expected clusters. Starting from an initial guess of the two distributions, EM iteratively calculates the probability of each point in the KLD-MDS embedding belonging to the two clusters and, using these probabilities as weights, re-estimating the mean and covariance matrix of the two distributions. This process is repeated until two distributions that represent the clustering in the coordinates are obtained. Points in the KLD-MDS embedding with $p > 0.01$ for a given distribution were then assigned to the cluster that distribution represents. We anticipate that this process should group datasets by the formulation of the sample if the subvisible particles in the two formulations are significantly different in morphology.

Effect of Formulation Differences

The two approved Omontys® formulations contain identical concentrations of peginesatide as well as sorbitol, but the SUV uses a phosphate buffer and contains micromolar concentrations of Tween 20, whereas the MUV uses a methionine buffer and contains phenol as a preservative. These differences in formulation might have affected particle formation mechanism(s) and consequently the resulting particle morphologies in the two formulations. However, it is also possible that these differences might have affected the capability of FIM to accurately detect and measure morphological features of particles. For example, differences in refractive index (RI) can alter the apparent transparency of particles in FIM[49]. It is therefore possible that different populations of particles detected using KLD-MDS analysis could reflect formulation differences, rather than differences in particle morphologies.

To test whether the KLD-MDS analysis was detecting only formulation differences between SUV and MUV formulations rather than differences in the morphologies of particles within the respective formulations, we created a standard population of protein aggregates, and then spiked small amounts of these pre-formed aggregates into solutions whose excipient concentrations matched those of the SUV and MUV formulations. FIM and KLD-MDS analyses were then conducted to determine whether the standard protein aggregate particles appeared to be different in the two formulations.

Preparation of suspensions of standard protein particles in MUV and SUV formulations

To generate a standard suspension of protein aggregates, 0.5 mg/mL IVIg in 1×PBS were centrifuged at 20,000 g at 4°C for 20 minutes to remove small aggregates. 0.9 mL aliquots of the supernatant were placed in 1.5 mL microcentrifuge tubes and exposed to six freeze-thaw cycles. Each cycle consisted of suspension in liquid nitrogen for 2 minutes followed by suspension in a hot water bath at 30 °C for 6 minutes. Two buffers were made which matched the excipient profile of the SUV and MUV formulation as described in the prescribing information. The SUV formulation contained 47 mg/mL sorbitol, 2.3 mg/mL anhydrous sodium phosphate monobasic, 0.12 mg/mL sodium phosphate dibasic, and 0.04 mg/mL polysorbate 20. The MUV formulation contained 47 mg/mL sorbitol, 5 mg/mL phenol, 1.5 mg/mL L-methionine, and 0.6 mg/mL glacial acetic acid. The pH of the SUV and MUV formulations were 6.0 and 5.4, respectively. 50 μL aliquots of the standard suspension fo protein aggregates were mixed with 950 μL aliquots of either the SUV and

MUV formulation buffers to form triplicate 1 mL aliquots of IVIg particles in the SUV and MUV formulations.

Analysis of standard protein particles spiked into MUV and SUV solutions

Flow-imaging was performed on three 300 μL aliquots of each sample using a FlowCam® VS (Fluid Imaging Technologies, Inc., Scarborough, ME) instrument using a 100-μm flow cell and the 10× objective. The flow cell was flushed with 1% Hellmanex III solution followed by water before running samples and with water between individual measurements. Images obtained from the FlowCam were analyzed using the KLD-MDS algorithm using the same particle properties used to analyze the full Omontys® dataset (i.e. size, aspect ratio, circularity, and average intensity). Since the FlowCam model used in this analysis collects RGB color images as opposed to the grayscale images available in the initial dataset collected by the FDA researchers, color images were converted to grayscale images prior to analysis. RGB pixel values were converted to grayscale intensities using the luminosity conversion:

$$c_{gray} = 0.2126c_{red} + 0.7152c_{green} + 0.0722c_{blue} \quad (10)$$

where c_{gray} is a grayscale pixel value (i.e. pixel intensity) and c_{red}, c_{green}, and c_{blue} are the red, green, and blue channel values, respectively, of the corresponding RGB pixel.

Results

Image Analysis

Figure 2 (a, b) shows representative histograms of the particle size distributions obtained from the image analysis for randomly chosen SUV and MUV vials, respectively, taken from experiment C. The two samples have relatively similar size distributions and would otherwise be difficult to differentiate by visual analysis, especially in a quantitative manner.

KLD-MDS Plots

Figure 3 (a, b, c) shows the two-dimensional embeddings obtained from the KLD-MDS analysis for experiments A, B, and C, respectively. Grouping of samples by formulation can be observed in all three figures. For all three experiments SUV samples appear closer to other SUV samples than MUV samples in the KLD-MDS embedding. Although three MUV samples overlap with SUV samples in figure 3 (b) and two samples are separated from the MUV samples in figure 3 (c), in general MUV samples also generally exhibit clustering behavior.

Clustering Analysis

Table 1 shows the average silhouette coefficient for each formulation in each of the three experiments as well as the average overall silhouette coefficient for the experiment. Experiments A and C have approximately similar values of the silhouette coefficient while Experiment B has a significantly lower average silhouette coefficient as was expected from Figure 3 (b). Note that, with the exception of MUV vials in experiment B, all formulations have moderate positive average silhouette coefficients, indicating that the datasets are generally clustered by formulation.

Figure 3 (a, b, c) also shows the boundary of clusters obtained from EM analysis for Experiments A, B, and C, respectively. The boundaries represent the $p > 0.01$ probability region of the two 2D Gaussians representing the two clusters. The clusters obtained from EM generally separate the datasets by formulation. As shown in figure 3 (a), 14 of the 16 datasets in Experiment A are correctly sorted by formulation. Similarly, 16 of the 19 datasets are correctly sorted by formulation for Experiment B and 21 of the 24 datasets are correctly sorted for Experiment C.

Figure 4 shows the two-dimensional embedding obtained from performing the KLD-MDS analysis using data from all three experiments at once. In spite of the differences in FlowCam analysis parameters used by the FDA for the three experiments, with proper normalization the KLD-MDS analysis still indicates a significant formulation dependency on the properties of subvisible particles present in the sample. Although only a single cluster appears in the figure, the coordinates are segregated by formulation within this cluster with the majority of SUV datasets appearing on the left of the cluster and the majority of MUV datasets appearing on the right.

Effect of Omontys formulations on FIM analysis

Figure 5 shows the results of KLD-MDS analysis of suspensions of standard protein aggregates in the MUV and SUV formulations. The populations of standard protein aggregates in the two formulations were indistinguishable by KLD-MDS analysis, with silhouette coefficients of 0.023 and 0.14 for the SUV and MUV formulations, respectively.

Discussion

Automated image processing techniques are continuously growing in prominence in performance in tasks ranging from object recognition[50] to developing self-driving cars[51]. These image processing techniques have also shown potential in areas of biomedical research such as diagnosing lung cancer from CT imaging[52] and diagnosing melanoma from smartphone images[53]. In these cases, image analysis techniques could diagnose cancers with accuracy meeting or exceeding that of trained clinicians. These image processing techniques may greatly enhance the amount of subvisible particle information that can be extracted from FIM. Although machine learning techniques have recently been used to analyze these images, these approaches have primarily been used to perform simple classification tasks such as the differentiation of silicone oil droplets from protein aggregates[54–56]. It is anticipated that more sophisticated image analysis techniques such as this Kullback-Leibler divergence-based approach may allow users to perform more complicated characterization tasks.

KLD-MDS can be applied to FIM datasets to identify a difference in particle morphology between the two formulations of peginesatide. In the 2D embedding obtained from KLD-MDS, most FIM measurements in the three experiments cluster with other measurements taken from the same formulation, e.g., FIM datasets from MUV samples are closest to datasets from other MUV samples, and similarly FIM datasets from SUV samples are closest to datasets from other SUV samples. The clustering is most apparent for datasets collected in Experiments A and C. Although the measurements taken in Experiment B exhibit less well-resolved clusters, the majority of datasets is still reasonably clustered in the KLD-MDS embedding. These formulation-dependent differences can also be identified when comparing FIM datasets from multiple experiments at once with proper normalization. Although the KLD-MDS embedding obtained from performing the analysis on the combined FIM data sets from all three experiments A, B and C lacks the more resolved clustering observed for the individual experiments, the new datasets still appear to be readily sorted by the formulation from which the dataset was obtained.

Clustering analysis by EM analysis was performed using a fixed number of clusters (two), as opposed to traditional clustering problems where the number of clusters may be a fitted variable. The clustering of these datasets as determined by application of EM algorithms agrees well with a simple visual analysis of the KLD-MDS plots, in which it is apparent that the data are largely segregated into two groups composed mostly of either SUV- or MUV-derived samples. Because the goal of this analysis is to see how well an unsupervised clustering technique can recover the anticipated formulation-dependent subvisible particle differences, fixing the number of clusters at two is sufficient for our purposes.

In the KLD-MDS analysis, the sets of data from the SUV and the MUV lots were relatively well-separated, so the value of p used to specify the decision boundary of the clusters calculated by EM had relatively little effect on which points were contained within the given clusters. However, we anticipate that the choice of p value will be more important for eventual applications of these techniques in process monitoring and control. For instance, consider a process-monitoring application in which FIM datasets recorded on new lots of a product are compared to sets of FIM measurements from older lots that are known to meet product specifications. p is the significance level threshold that determines whether or not a given lot has particles that match those found in the lots that meet product specifications. In this example p should be set high enough to identify samples with particles that dramatically vary in subvisible particle populations and potential immunogenicity from the normal product, but low enough to avoid unnecessary process downtime due to an incorrectly identified process upset. Optimization of the value of p needed to balance these two risks was outside the scope of this study.

The clustering of datasets by formulation suggests that populations of subvisible particles in the two peginesatide formulations exhibit significantly different morphologies; these differences in particle characteristics are larger than any differences between samples of the same formulation. These subtle differences in particle morphology are difficult to detect, even at a qualitative level, using currently standard FIM analyses like the particle size distribution histograms shown in Figure 1. Although histograms such as those shown in Figure 1 are visually difficult to differentiate even in a qualitative manner, the KLD-based approach is capable of identifying a quantitative difference in subvisible particle populations between the two formulations. This variation correlates with the difference in the frequency of severe ADRs that was observed in the clinic, suggesting that a change in particle morphology could have contributed to the change in immunogenicity.

The SUV and MUV formulations of peginesatide were slightly different in composition and were filled into different containers. These differences in formulation and container-closure systems likely contributed to differences in the mechanisms by which the particles were created, and in turn resulted in differences in particle morphology that could be detected by our KLD-MDS analyses of FIM datasets. An alternative explanation is that apparent differences in particle populations detected in SUV and MUV datasets are the result of formulation-generated biases in the FIM analyses, rather than actual differences in particle morphology. This alternative explanation can be discounted, because an analysis conducted on suspensions of standard aggregates in the same formulations yielded no detectable differences in particle populations that could be ascribed to formulation effects on FIM images.

Although the FDA researchers had previously identified differences in the numbers of subvisible particles found in the clinical and marketed Omontys® formulations[44], we stress that we have identified a separate difference in between their populations of subvisible particles. The FDA reported that the MUV formulation generally had higher concentrations of subvisible particles than did the SUV formulation. In contrast, our analysis is not influenced by the concentration of particles in the sample but is instead focused on identifying differences in morphology between *populations* of particles in the two samples. Our analysis therefore identifies particle morphology as a separate factor that could have contributed to the ADRs to the marketed Omontys® formulation.

Our findings in conjunction with the earlier findings of the FDA[44] indicate that the two Omontys® formulations exhibited substantial differences in subvisible particle populations. Although both formulations met particle concentration limitations set by USP <788> for particles of size larger than 10μm and larger than 25 μm[44], subvisible particle populations for the two formulations differed both in particle concentration and, as is evident from our KLD-MDS analysis, particle morphological properties. Although they are associated with the serious ADRs experienced by patients receiving the MUV formulations of Omontys®, neither the population distributions of subvisible particles nor their respective morphologies can be causally linked to the ADRs. It is apparent in retrospect, however, that had the differences in concentrations and morphologies of subvisible particles between the safe, clinically-tested SUV formulation and the ADR-provoking MUV formulation been known, a red flag should have been raised prior to initiation of marketing.

Acknowledgments

This study was funded by the National Institutes of Health, grant RO1 EB006006.

References

1. Walsh G. Biopharmaceutical benchmarks 2014. Nat Biotechnol. 2014; 32(10):992–1000. [PubMed: 25299917]
2. Schellekens H. Immunogenicity of therapeutic proteins: Clinical implications and future prospects. Clinical Therapeutics. 2002; 24(11):1720–1740. [PubMed: 12501870]
3. Schellekens H. Immunogenicity of therapeutic proteins. Nephrology Dialysis Transplantation. 2003; 18(7):1257–1259.

4. Tamilvanan S, Raja NL, Sa B, Basu SK. Clinical concerns of immunogenicity produced at cellular levels by biopharmaceuticals following their parenteral administration into human body. JOURNAL OF DRUG TARGETING. 2010; 18(7):489–498. [PubMed: 20192653]
5. Singh SK. Impact of product-related factors on immunogenicity of biotherapeutics. J. Pharm. Sci. 2011; 100(2):354–387. [PubMed: 20740683]
6. Hartung HP, Munschauer F, Schellekens H. Significance of neutralizing antibodies to interferon beta during treatment of multiple sclerosis: expert opinions based on the Proceedings of an International Consensus Conference. European Journal of Neurology. 2005; 12(8):588–601. [PubMed: 16053466]
7. Goodin DS, Frohman EM, Hurwitz B, O'Connor PW, Oger JJ, Reder AT, Stevens JC. Neutralizing antibodies to interferon beta: Assessment of their clinical and radiographic impact: An evidence report - Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology. 2007; 68(13):977–984. [PubMed: 17389300]
8. Malucchi S, Sala A, Gilli F, Bottero R, Di Sapio A, Capobianco M, Bertolotto A. Neutralizing antibodies reduce the efficacy of beta IFN during treatment of multiple sclerosis. Neurology. 2004; 62(11):2031–2037. [PubMed: 15184610]
9. Ridker PM, Tardif JC, Amarenco P, Duggan W, Glynn RJ, Jukema JW, Kastelein JJP, Kim AM, Koenig W, Nissen S, Revkin J, Rose LM, Santos RD, Schwartz PF, Shear CL, Yunis C, Investigators S. Lipid-Reduction Variability and Antidrug-Antibody Formation with Bococizumab. N Engl J Med. 2017; 376(16):1517–1526. [PubMed: 28304227]
10. Schellekens H. Bioequivalence and the immunogenicity of biopharmaceuticals. Nature Reviews Drug Discovery. 2002; 1(6):457–462. [PubMed: 12119747]
11. Hermeling S, Crommelin DJA, Schellekens H, Jiskoot W. Structure-immunogenicity relationships of therapeutic proteins. Pharmaceutical Research. 2004; 21(6):897–903. [PubMed: 15212151]
12. Sauerborn M, Brinks V, Jiskoot W, Schellekens H. Immunological mechanism underlying the immune response to recombinant human protein therapeutics. TRENDS IN PHARMACOLOGICAL SCIENCES. 2010; 31(2):53–59. [PubMed: 19963283]
13. van Beers MM, Jiskoot W, Schellekens H. On the role of aggregates in the immunogenicity of recombinant human interferon beta in patients with multiple sclerosis. J Interferon Cytokine Res. 2010; 30(10):767–75. [PubMed: 20874254]
14. van Beers MM, Bardor M. Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. Biotechnol J. 2012; 7(12):1473–84. [PubMed: 23027660]
15. Torosantucci R, Schöneich C, Jiskoot W. Oxidation of therapeutic proteins and peptides: structural and biological consequences. Pharm Res. 2014; 31(3):541–53. [PubMed: 24065593]
16. Ratanji KD, Derrick JP, Dearman RJ, Kimber I. Immunogenicity of therapeutic proteins: influence of aggregation. J Immunotoxicol. 2014; 11(2):99–109. [PubMed: 23919460]
17. Wang W, Singh SK, Li N, Toler MR, King KR, Nema S. Immunogenicity of protein aggregates--concerns and realities. Int J Pharm. 2012; 431(1–2):1–11. [PubMed: 22546296]
18. Chisholm CF, Baker AE, Soucie KR, Torres RM, Carpenter JF, Randolph TW. Silicone Oil Microdroplets Can Induce Antibody Responses Against Recombinant Murine Growth Hormone in Mice. J Pharm Sci. 2016; 105(5):1623–32. [PubMed: 27020987]
19. Chisholm CF, Soucie KR, Song JS, Strauch P, Torres RM, Carpenter JF, Ragheb JA, Randolph TW. Immunogenicity of Structurally Perturbed Hen Egg Lysozyme Adsorbed to Silicone Oil Microdroplets in Wild-Type and Transgenic Mouse Models. J Pharm Sci. 2017; 106(6):1519–1527. [PubMed: 28216023]
20. Fradkin AH, Carpenter JF, Randolph TW. Immunogenicity of Aggregates of Recombinant Human Growth Hormone in Mouse Models. JOURNAL OF PHARMACEUTICAL SCIENCES. 2009; 98(9):3247–3264. [PubMed: 19569057]
21. Fradkin AH, Carpenter JF, Randolph TW. Glass Particles as an Adjuvant: A Model for Adverse Immunogenicity of Therapeutic Proteins. Journal of Pharmaceutical Sciences. 2011; 100(11): 4953–4964. [PubMed: 21721003]
22. Van Beers MMC, Gilli F, Schellekens H, Randolph TW, Jiskoot W. Immunogenicity of recombinant human interferon beta interacting with particles of glass, metal, and polystyrene. Journal of Pharmaceutical Sciences. 2012; 101(1):187–199. [PubMed: 21918983]

23. Shomali M, Freitag A, Engert J, Siedler M, Kaymakcalan Z, Winter G, Carpenter JF, Randolph TW. Antibody responses in mice to particles formed from adsorption of a murine monoclonal antibody onto glass microparticles. J Pharm Sci. 2014; 103(1):78–89. [PubMed: 24227137]
24. Freitag AJ, Shomali M, Michalakis S, Biel M, Siedler M, Kaymakcalan Z, Carpenter JF, Randolph TW, Winter G, Engert J. Investigation of the Immunogenicity of Different Types of Aggregates of a Murine Monoclonal Antibody in Mice. Pharmaceutical Research. 2015; 32(2):430–444. [PubMed: 25123991]
25. Shomali M, Tanriverdi S, Freitag AJ, Engert J, Winter G, Siedler M, Kaymakcalan Z, Carpenter JF, Randolph TW. Dose Levels in Particulate-Containing Formulations Impact Anti-drug Antibody Responses to Murine Monoclonal Antibody in Mice. Journal of Pharmaceutical Sciences. 2015; 104(5):1610–1621. [PubMed: 25737325]
26. Seefeldt MB, Rosendahl MS, Cleland JL, Hesterberg LK. Application of high hydrostatic pressure to dissociate aggregates and refold proteins. Curr Pharm Biotechnol. 2009; 10(4):447–55. [PubMed: 19519422]
27. den Engelsman J, Garidel P, Smulders R, Koll H, Smith B, Bassarab S, Seidl A, Hainzl O, Jiskoot W. Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. Pharm Res. 2011; 28(4):920–33. [PubMed: 20972611]
28. Rosenberg AS. Effects of Protein Aggregates: An Immunological Perspective. AAPS Journal. 2006; 8(3):E501–E507. [PubMed: 17025268]
29. Jiskoot W, Kijanka G, Randolph TW, Carpenter JF, Koulov AV, Mahler HC, Joubert MK, Jawa V, Narhi LO. Mouse Models for Assessing Protein Immunogenicity: Lessons and Challenges. J Pharm Sci. 2016; 105(5):1567–75. [PubMed: 27044944]
30. Ryff JC, Schellekens H. Immunogenicity of rDNA-derived pharmaceuticals. Trends Pharmacol Sci. 2002; 23(6):254–256. [PubMed: 12084625]
31. Hochuli E. Interferon immunogenicity: Technical evaluation of interferon- alpha 2a. J Interferon Cytokine Res. 1997; 17:S15–S21. [PubMed: 9241611]
32. Prummer O. Endogenous antibodies directed against interferon-alpha: paradigm of the immunogenicity of cytokines and its clinical impact. Med Welt. 1998; 49(6):267–274.
33. Ring J, Stephan W, Brendel W. Anaphylactoid Reactions to Infusions of Plasma-Protein and Human-Serum Albumin - Role of Aggregated Proteins and of Stabilizers Added during Production. Clin Allergy. 1979; 9(1):89–97. [PubMed: 421339]
34. Moore WV, Leppert P. Role of Aggregated Human Growth-Hormone (Hgh) in Development of Antibodies to Hgh. J Clin Endocrinol Metab. 1980; 51(4):691–697. [PubMed: 7419661]
35. Ahmadi M, Bryson CJ, Cloake EA, Welch K, Filipe V, Romeijn S, Hawe A, Jiskoot W, Baker MP, Fogg MH. Small amounts of sub-visible aggregates enhance the immunogenic potential of monoclonal antibody therapeutics. Pharm Res. 2015; 32(4):1383–94. [PubMed: 25319104]
36. Joubert MK, Hokom M, Eakin C, Zhou L, Deshpande M, Baker MP, Goletz TJ, Kerwin BA, Chirmule N, Narhi LO, Jawa V. Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses. J Biol Chem. 2012; 287(30):25266–79. [PubMed: 22584577]
37. Rombach-Riegraf V, Karle AC, Wolf B, Sordé L, Koepke S, Gottlieb S, Krieg J, Djidja MC, Baban A, Spindeldreher S, Koulov AV, Kiessling A. Aggregation of human recombinant monoclonal antibodies influences the capacity of dendritic cells to stimulate adaptive T-cell responses in vitro. PLoS One. 2014; 9(1):e86322. [PubMed: 24466023]
38. Barandun S, Jeunet F, Kistler P, Isliker H. Intravenous Administration of Human Gamma-Globulin. Vox Sanguinis. 1962; 7(2):157–74. [PubMed: 13864762]
39. Ellis EF, Henney CS. Adverse Reactions Following Administration of Human Gamma Globulin. Journal of Allergy. 1969; 43(1):45–54. [PubMed: 4178567]
40. Joubert MK, Luo QZ, Nashed-Samuel Y, Wypych J, Narhi LO. Classification and Characterization of Therapeutic Antibody Aggregates. Journal of Biological Chemistry. 2011; 286(28):25118–25133. [PubMed: 21454532]
41. Bessa J, Boeckle S, Beck H, Buckel T, Schlicht S, Ebeling M, Kiialainen A, Koulov A, Boll B, Weiser T, Singer T, Rolink AG, Iglesias A. The immunogenicity of antibody aggregates in a novel transgenic mouse model. Pharm Res. 2015; 32(7):2344–59. [PubMed: 25630815]

42. Freitag AJ, Shomali M, Michalakis S, Biel M, Siedler M, Kaymakcalan Z, Carpenter JF, Randolph TW, Winter G, Engert J. Investigation of the Immunogenicity of Different Types of Aggregates of a Murine Monoclonal Antibody in Mice. Pharm Res. 2014; 32(2):430–44. [PubMed: 25123991]
43. Maddux NR, Daniels AL, Randolph TW. Microflow Imaging Analyses Reflect Mechanisms of Aggregate Formation: Comparing Protein Particle Data Sets Using the Kullback-Leibler Divergence. J Pharm Sci. 2017; 106(5):1239–1248. [PubMed: 28159641]
44. Kotarek J, Stuart C, De Paoli SH, Simak J, Lin TL, Gao Y, Ovanesov M, Liang Y, Scott D, Brown J, Bai Y, Metcalfe DD, Marszal E, Ragheb JA. Subvisible Particle Content, Formulation, and Dose of an Erythropoietin Peptide Mimetic Product Are Associated With Severe Adverse Postmarketing Events. J Pharm Sci. 2016; 105(3):1023–7. [PubMed: 26886324]
45. FDA. Highlights Of Perscribing Information: Omontys. 2012
46. Kalonia C, Kumru OS, Prajapati I, Mathaes R, Engert J, Zhou S, Middaugh CR, Volkin DB. Caclulating the Mass of Subvisible Protein Particles with Improved Accuracy Using Microflow Imaging Data. J Pharm Sci. 2015; 104(2):536–47. [PubMed: 25302696]
47. Rousseeuw PJ. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. Journal of Computational and Applied Mathematics. 1987; 20:53–65.
48. Dempster A, Laird N, Rubin DB. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. 1977; 39(1):1–38.
49. Zölls S, Gregoritza M, Tantipolphan R, Wiggenhorn M, Winter G, Friess W, Hawe A. How Subvisible Particles Become Invisible — Relevance of the Refractive Index for Protein Particle Analysis. Pharm. Biotechnol. 2013; 102(5):1434–46.
50. Szegedy S, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going Deeper with Convolutions. IEEE. 2015
51. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J. End to End Learning for Self-Driving Cars. arXiv: 160407316. 2016. http://arxiv.org/abs/1604.07316
52. Arindra A, Setio A, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ, Winglker Wille MM, Naqibullah M, Sánchez CI, van Ginneken B. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. IEEE Trans Med Imaging. 2016; 35(5):1160–9. [PubMed: 26955024]
53. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017; 542:115–8. [PubMed: 28117445]
54. Strehl R, Rombach-Riegraf V, Diez M, Egodage K, Bluemel M, Jeschke M, Koulov AV. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: A novel multiparametric image filter for sub-visible particles in microflow imaging analysis. Pharm. Res. 2012; 29(2):594–602. [PubMed: 21948455]
55. Zölls S, Weinbuch D, Wiggenhorn M, Winter G, Friess W, Jiskoot W, Hawe A. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. AAPS J. 2013; 15(4):1200–11. [PubMed: 23996547]
56. Saggu M, Patel AR, Koulis T. A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy. Pharm. Res. 2017; 34(2):479–91. [PubMed: 27995522]

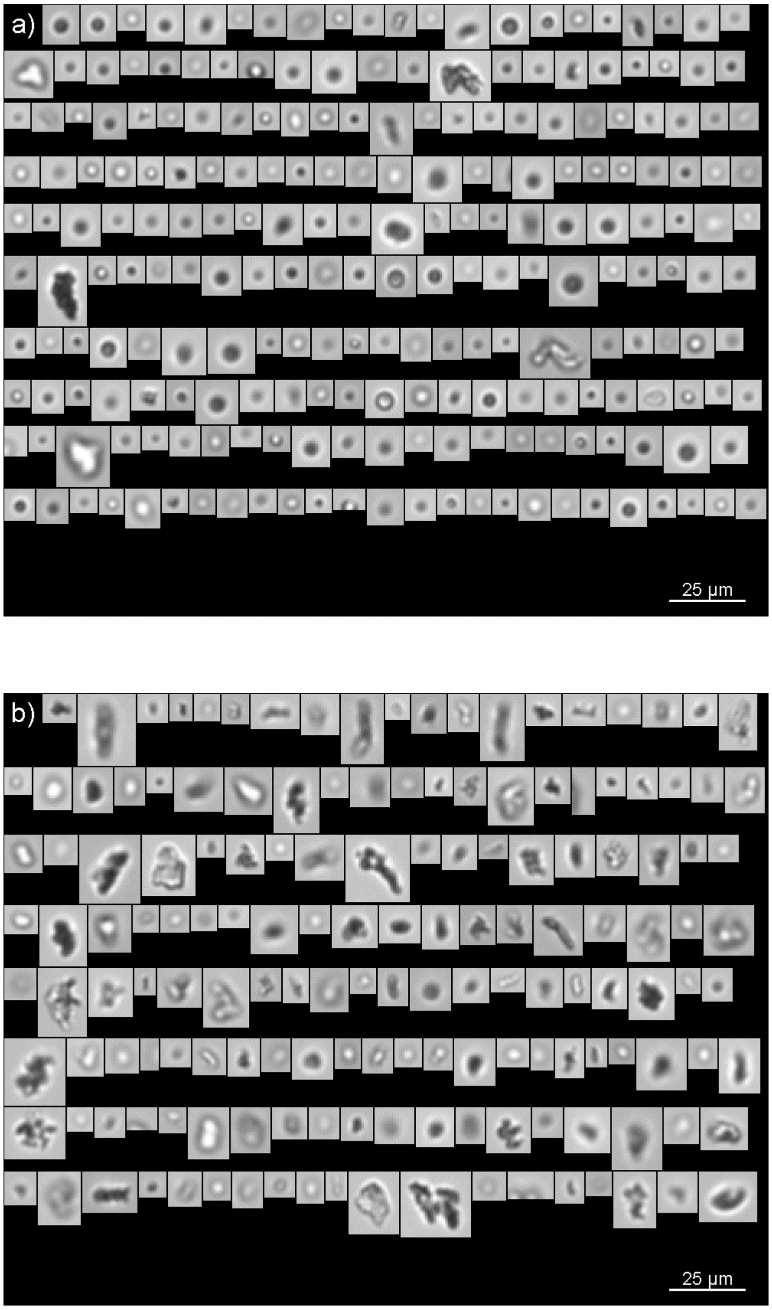

Figure 1.
Sample images taken from a) SUV and b) MUV samples analyzed in experiment C.

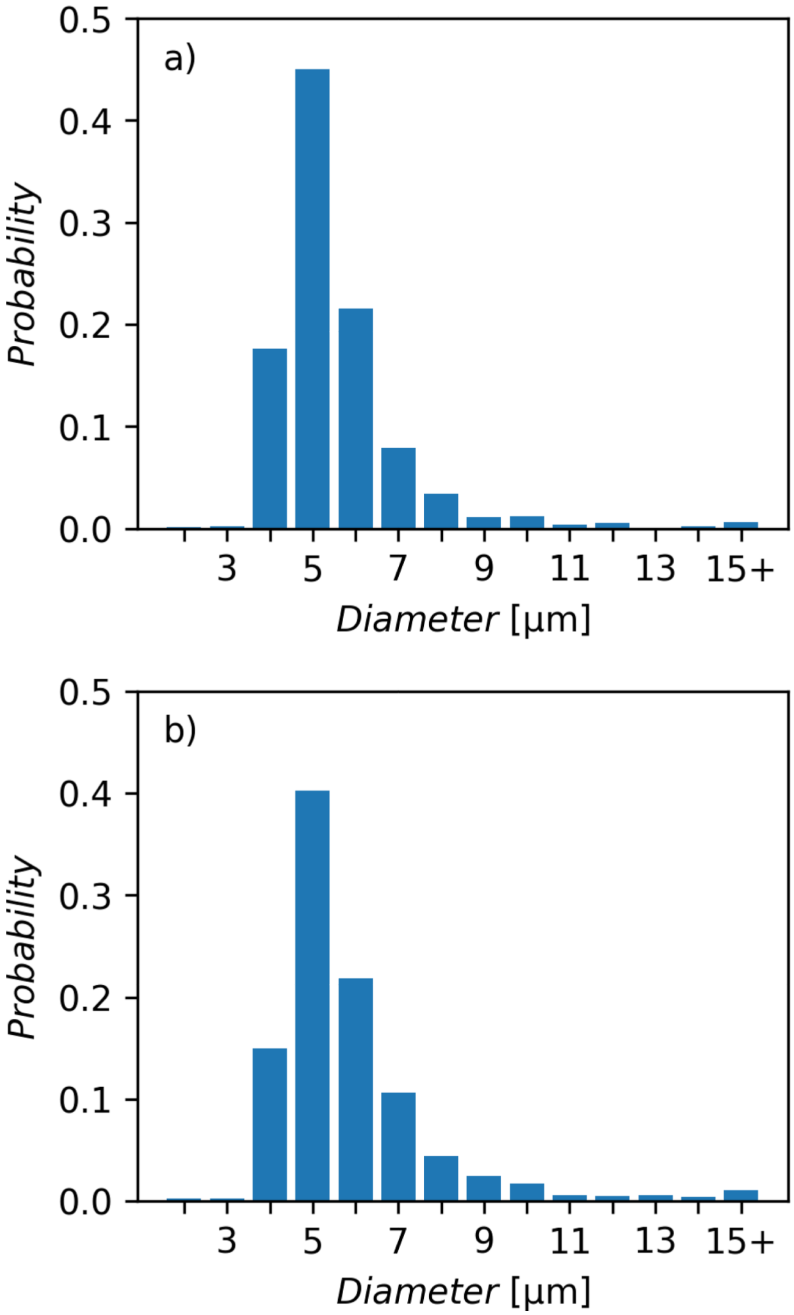

Figure 2.

Histograms of particle size of a random a) SUV and b) MUV sample taken from experiment C. The MUV sample shows the presence of generally larger particles than the SUV sample. However, the histograms otherwise indicate generally similar distributions of particle sizes.

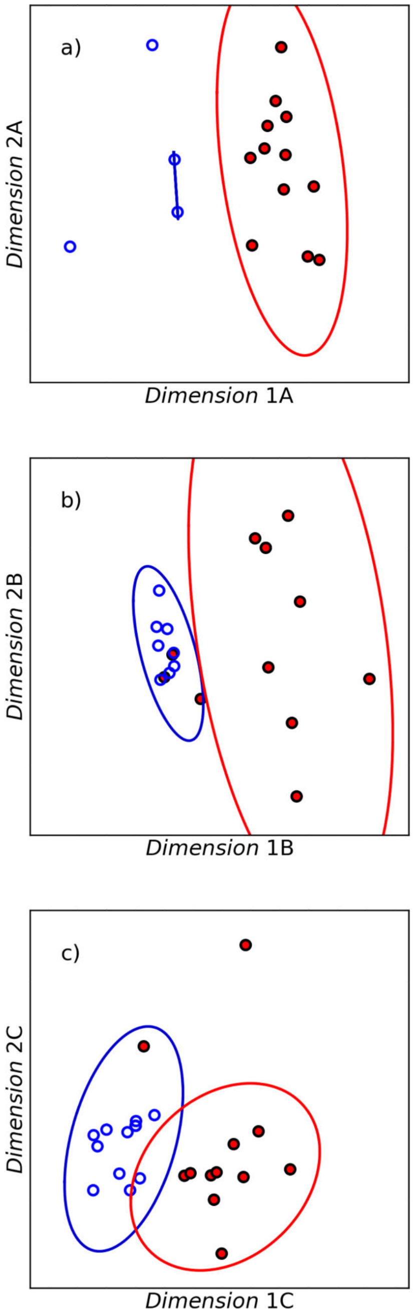

Figure 3.

KLD-MDS plots of the FIM data collected in (a) experiment A, (b) experiment B, and (c) experiment C. Plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Also shown on these figures are the $p > 0.01$ regions of the 2D Gaussians obtained from expectation maximization as represented by the red and blue ovals. As can be seen visually, datasets in experiment B exhibit moderate clustering by formulation and datasets in experiments A and C exhibit much more substantial clustering by formulation. This clustering is confirmed by EM; most SUV datasets (open blue circles) are within the $p > 0.01$ region of the SUV cluster (blue oval) and most MUV datasets (red-filled circles) are within the $p > 0.01$ region of the MUV cluster (red oval).

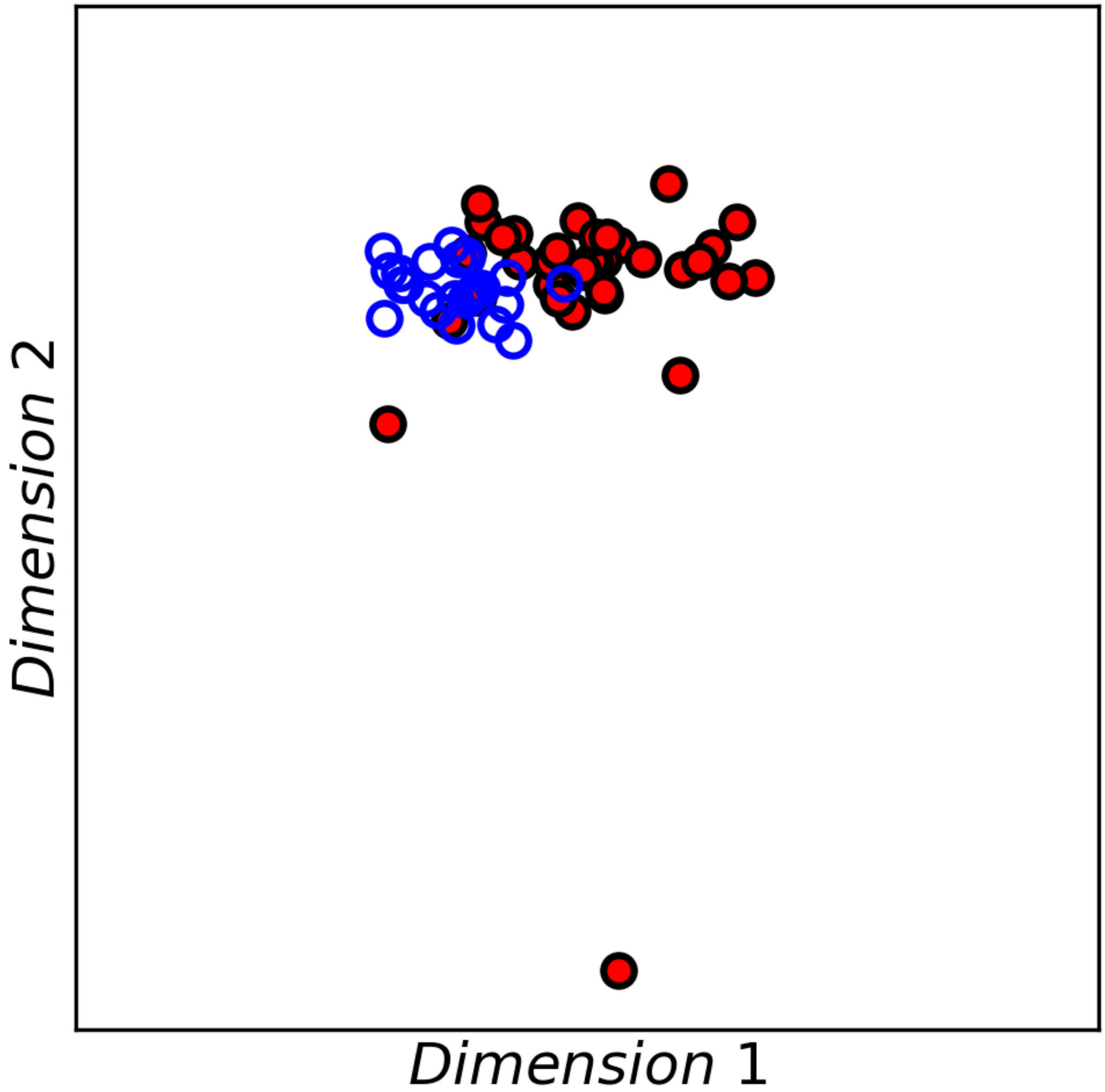

Figure 4.

KLD-MDS plot showing data for all three experiments. Plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Although the plot lacks resolved clusters for the SUV (open blue circles) and MUV samples (red-filled circles), the datasets still exhibit noticeable segregation by formulation indicative of a formulation dependency in the subvisible particle populations.

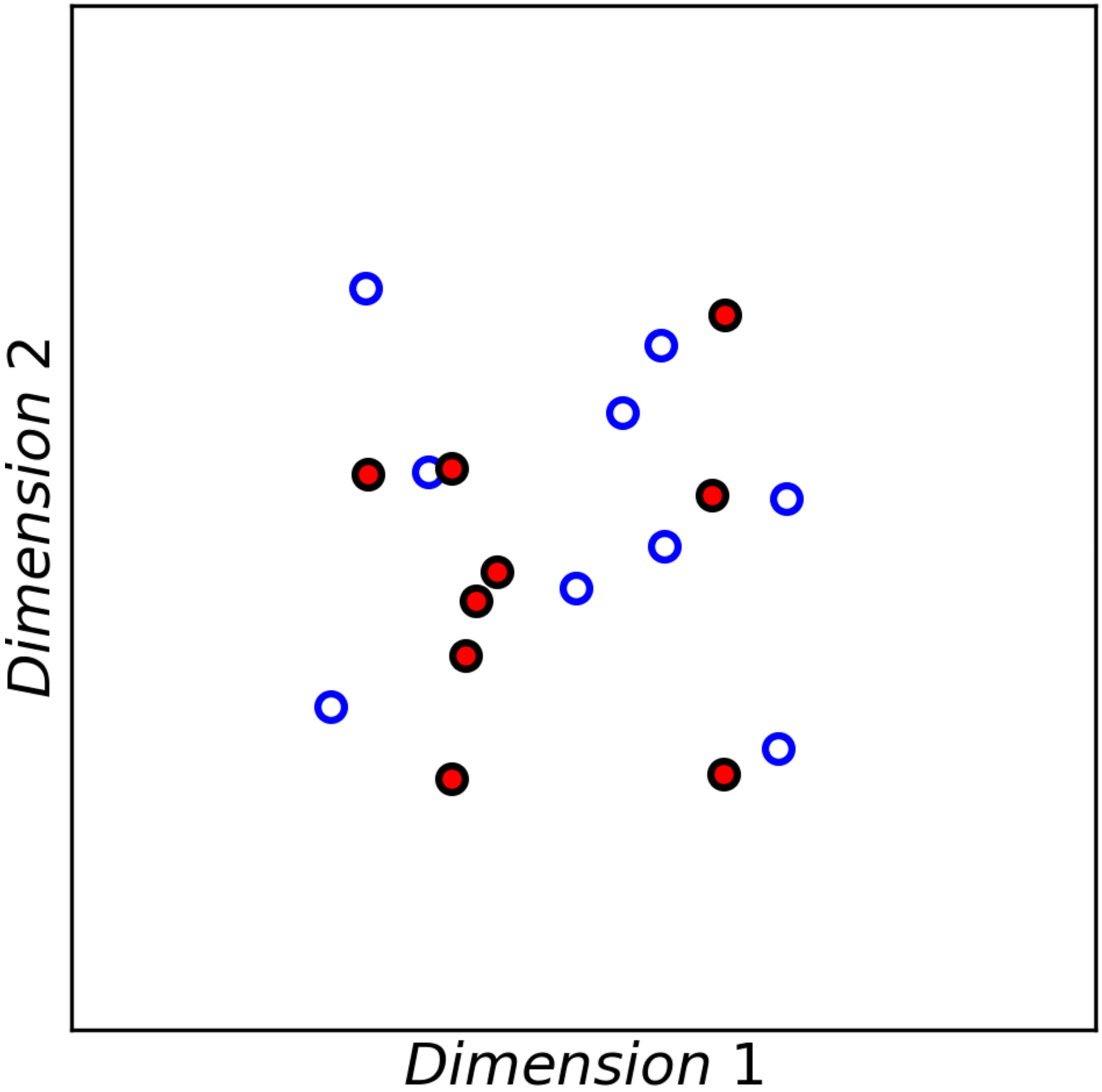

Figure 5.

KLD-MDS plot of the IVIg particles suspended in SUV (red-filled circles) and MUV (open blue circles) formulation buffers. As in figure 3 plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Visually the clustering present in figures 3–4 is mostly absent when both formulations contain identical particles, suggesting that the refractive indices of the two formulations does not significantly influence the KLD-MDS analysis.

Table 1

Silhouette coefficients for each of the three experiments.

Experiment	SUV Silhouette Coefficient	MUV Silhouette Coefficient	Average Silhouette Coefficient
A	0.37	0.53	0.49
B	0.78	−0.04	0.31
C	0.71	0.25	0.48