



Assaying Protein Quality Using Single Molecule Detection

Presented at the 2009 ESACT meeting, Dublin, Ireland

David G. Whitehead and Jean-François Hamel*

MIT, Dept. of Chemical Engineering, Cambridge, MA 02139

*Email: jhamel@mit.edu

Introduction: Biopharmaceuticals rely heavily on protein production to create drugs and other medicinal products. The protein produced needs to be of high quality (i.e., properly folded and not aggregated) to be used as an effective drug, because protein aggregates can affect the immunogenicity and effectiveness of the drug. Many companies create multiple clone lines of cells and measure the quality of protein being produced. They focus on factors such as quantity and quality of protein production to decide which lines of cells to continue growing. A simple and quick assay of the quality of protein produced is therefore of great importance to the biopharmaceutical industry.

Using the ConSense Analyzer—fluIT's single molecule detection technology, accurate stochastic fluorescence spectroscopy—it is possible to differentiate between monomer and aggregates in a sample. This can be achieved in a quick process using very small quantities of culture. The fluIT software Virtual Lab can analyze the sample and graph the distribution of monomer and aggregate in a sample. Further analysis (e.g., using Excel) can help quantify a more exact ratio of monomer to aggregate in the sample.

The protein used in this study was a lymphotoxin-β receptor protein fused with IgG (LTβR-IgG). This protein is about 100 kDa and has 23 cystine bonds. When misfolded, LTβR-IgG is very prone to form large, complex aggregates. The main goals of this study were to setup a simple protein assay, minimizing the number of pre-analysis steps, and to evaluate the ability of the ConSense Analyzer to discriminate between monomeric and aggregated proteins and the reliability of the Analyzer data to measure the (known) ratio of monomer to aggregate in solution.

Materials: The ConSense Analyzer and glass chips with 14 wells, Nile Red dye (Invitrogen), and training were provided by fluIT Biosystems. Nile Red was prepared in a solution of 50% water and 50% ethanol, yielding a final concentration of 20 μmol Nile Red in solution. Biogen Idec provided samples of fresh cell-culture medium ("blank"), used cell-culture medium ("spent"), and purified at 0.5 mg/mL samples of monomer and aggregate (with unknown composition of multimers) LTβR-IgG fusion protein, produced by Chinese Hamster Ovary (CHO) cells.

Method: Sample Preparation (PBS) – Samples for analysis were made by adding 47.5 μL of 1× phosphate buffered saline buffer (PBS), 47.5 μL of monomer or aggregate LTβR-IgG, and 5 μL of Nile Red solution (so the concentration of Nile Red in the PBS samples was 1 μmol). The ConSense Analyzer recorded 30 s of fluorescence intensity from a 3 μL sample on a glass chip. With all 14 wells filled—with samples from the same solution, to establish replicable data—a chip took 5–10 min to analyze. First, samples with 100% monomer and 100% aggregate were compared using the Aggregation Analysis package (Figure 1). Then solutions were made with ratios of monomer to aggregate ranging from 100% monomer to 100% aggregate, in 10% increments. Data from these samples were used to create a protein-composition trend line.

Sample Preparation (blank) – Same as PBS, using blank (cell-free) medium from Biogen Idec in place of LTβR-IgG.

Sample Preparation (spent) – Spent medium had been prepared by Biogen Idec (according to the protocol developed by David Liu) and protein A was added to remove any protein in the supernatant, all before the medium arrived at the lab. Initially, samples were prepared the same as PBS, using spent medium in place of LTβR-IgG. Data from these samples included high levels of background noise. Subsequent samples were prepared by first running the spent medium through a 0.2 μm filter to remove any cell debris, and then preparing it the same as PBS.

Analysis – Two methods were used to analyze samples. The first method focused on the Molecular Brightness (MB) calculated by the ConSense Software Virtual Lab analysis package. Data from a blank sample and from the incremental monomer/aggregate ratio samples were put into a table of frequency of occurrence based on MB. The differences in frequency of brightness between the blank and the incremental samples were then charted.

The second method focused solely on the base brightness measured by the ConSense's photon detector. The brightness was then put into a frequency of occurrence table. A Gaussian curve using the equation: $e^{-\frac{(x-b)^2}{c}}$, where b and c are variable parameters, was fit to the curve of each protein-ratio sample (not shown). The values of b and c , along with how well the curve fit the data, were analyzed.

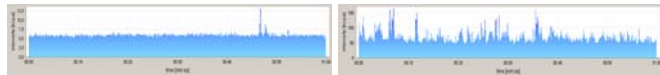


Figure 1: Fluorescence intensity of pure monomer (above, left; intensity scale to 12.5 kcps) and pure aggregate (above, right; intensity scale to 150 kcps) in PBS buffer, both at a concentration of 0.25 mg/mL. Molecular brightness of both samples are compared in the scatter plot (below). The aggregates (red) and monomers (blue) are clearly distinguishable.

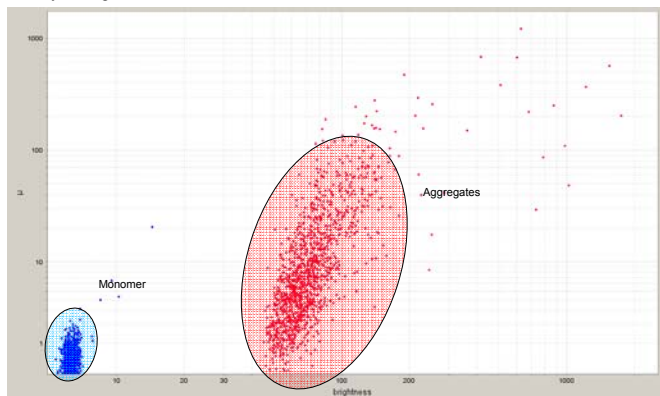


Figure 2: Molecular Brightness Analysis

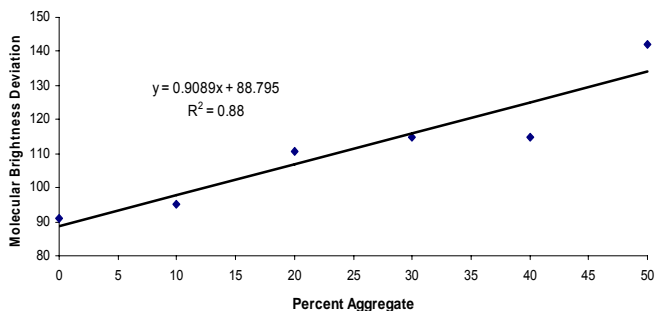


Figure 2: Average molecular brightness of ten-sample runs of 0%–50% aggregate LTβR-IgG samples, analyzed using the first method. The linear trend line fits well, with only the 40% and 50% aggregate data points having a notable deviation from the fit line.

Figure 3: Brightness Analysis

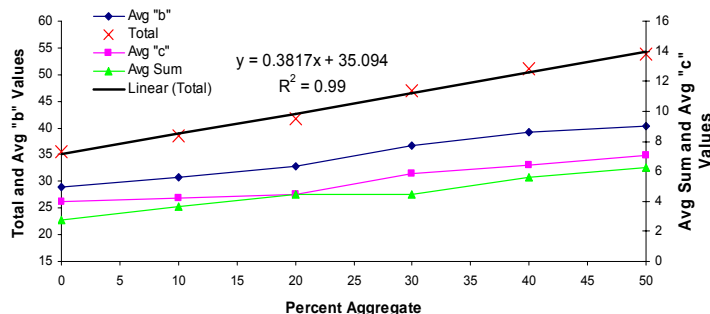


Figure 3: Average Gaussian-curve parameters and fit data from 29-sample runs of 0%–50% aggregate LTβR-IgG samples, analyzed using the second method focusing on brightness. The b , c , and fit values behave similarly, and when added together yielded a linear trend that fit the data well.

Results and Conclusions: Based on the first method of analysis, the blank yielded the lowest calculated MB. Increasing the aggregate concentration yielded higher calculated MB. When MB values were normalized by subtracting the frequency values of the blank from those of the other samples, the 100% monomer had the lowest value and the 100% aggregate had the highest value. These MB deviations for samples with 0%–50% aggregate are shown in Figure 2. In future assays, a linear trend line fit to the data could be used as a correlation to determine the unknown ratio of monomer to aggregate in a sample.

A similar aggregate-concentration based trend was found with the second method of analysis. The Gaussian equation variables b and c were optimized using the Excel solver function, based on attaining the best fit of the Gaussian curve to brightness data. The overall fit of the Gaussian to the data (the fit value) was measured by evaluating the absolute value of the deviation of each data point from the fit curve. The fit values and the values of b and c increased with increasing ratio of aggregate in the sample. An overall trend line was formed by adding the values of b and c and the fit value. These data for samples with 0%–50% aggregate are shown in Figure 3.

For the spent sample, the addition of a filtration step after centrifugation, background noise was noticeably reduced, increasing the statistical reliability of the findings. Statistical reliability may be improved by implementing effective protein purification steps before analysis.

One aspect of this assay that still needs some development is the way proteins are labeled. The dye used in this study, Nile Red, is a non-specific dye that binds to hydrophobic portions of molecules. This can result in the dye binding to molecules that are not of interest, increasing background noise. This causes a low signal-to-noise ratio that can make analysis difficult. Using specific binders—such as mAb or fAb, with a covalently bound dye label—should increase the signal-to-noise ratio as well as the specificity of the assay.

Acknowledgements: Thanks to: Biogen Idec for supplying samples of blank and spent media and LTβR-IgG protein and fluIT; Biosystems for providing the use of the ConSense Instrument and the Virtual Lab analysis software, which comes with the instrument; Sebastian Giehring and Thomas Tonn (fluIT) for training on how to use the ConSense and analysis software and assistance with data analysis; Graham Farrington, William Sisk, and Nels Pederson (Biogen Idec) for discussions of experiment development and data analysis; Scott Barksdale (Genesis Technologies) for introducing the ConSense technology to MIT and facilitating the team project; and Jamie Tyrrell (MIT) for textual support.

Key resources:

Schellekens H. Factors influencing the immunogenicity of therapeutic proteins. *Nephrol Dial Transplant* 2005; 20: vi3–vi9.
Hermeling S, Crommelin DJ, Schellekens H, Jiskoot W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res* 2004; 21: 897–903.
Robbins DC, Hirshman M, Wardzala LJ, Horton ES. High-molecular weight aggregates of therapeutic insulin: in vitro measurements of receptor binding and bioactivity. *Diabetes* 1988; 37: 56–59.