Multi-variate Data Analysis for Protein A Chromatography Column

Ketki Behere[a], Rachel Wollacott[b], Dr. Seongkyu Yoon[a]

[a]Dept. of Chemical Engineering, Massachusetts Bio Manufacturing Center, University of Massachusetts, Lowell, MA

[b]Process Development, MassBiologics of the University of Massachusetts Medical School, Boston, MA

3rd Biopharmaceutical Process & Quality Consortium (BPQC) Summit 2014

ABSTRACT

The objective of this project was to build a regression model which could explain the variability in the batch column chromatography dataset and provide some predictability to the user. The first step was to build a Batch Evolution Model (BEM) which could explain each chromatography step in detail and provide predictability power to the loading and elution step. The eventual goal of this project was to give a direct comparison between the different batches which had different resin. It was also required to identify the important response variables out of the given seven (Column height, column diameter, Asymmetry, HETP, resin capacity, loading and elution concentration) which provided significant contribution to the given dataset and thus can be extrapolated to batch chromatography in general. A validation of the BLM model was required to prove the efficiency of the model. The BEM model gave considerable predictability to the loading and elution steps and confirmed that those process steps were most important in the entire batch run. The BLM model gave definite segregation between the batches with different resin. Three response variables were identified to show significant contribution to the dataset. Those were the column height, HETP and elution peak concentration.

OBJECTIVE

✓ Build a regression model to explain the variability in the batch column chromatography dataset

✓ Provide predictability to the user

✓ Identify response variables with strong co-relation

The information explosion is the effect of abundance of data. As the amount of available data grows, the problem of managing the information becomes more difficult, which can lead to information overload.

Know your Data

Chromatography is a separation process of partitioning a protein between an insoluble stationary phase and a mobile phase. The UV sensor measurement at the beginning and end of loading and elution steps was considered to be critical to the quantification of protein.

Protein A Chromatography
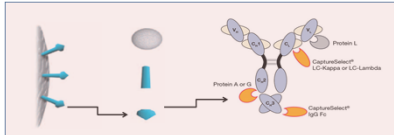
- Protein A is 42 kDa protein derived from a strain of *Staphylococcus aureus*.
- Affinity binding specific to Mab provides > 95% purity
- Consists of six regions, five of which bind IgG.
- Leachable Protein A needs to be removed
- Low pH elution may damage product
- Resin Shelf life needs to be assessed thoroughly.

METHODOLOGY

|	
Scaling	• Column volume was used as a reference
Segregation	• Batch was divided into phases as per Chromatography steps
Analysis	• BEM and BLM model developed using SIMCA 13.0
Validation	• Permutation tool in SIMCA was used to confirm the analysis

MODEL DEVELOPMENT

BATCH EVOLUTION MODEL (BEM)

Build a Batch Evolution Model (BEM) to explain each chromatography step. Categorize process steps into Phases:
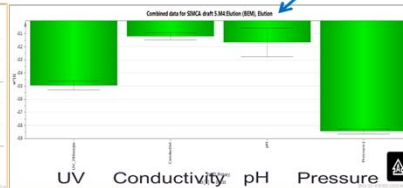
1. Equilibration
2. Loading
3. Wash
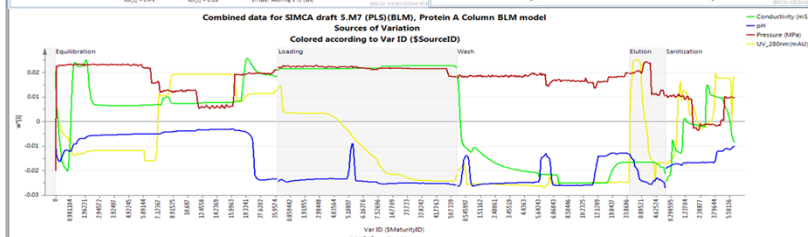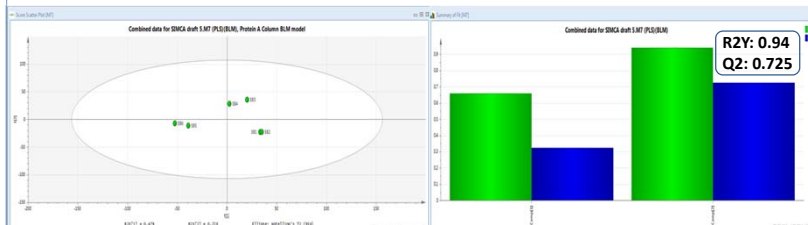4. Elution
5. Sanitization
6. Storage
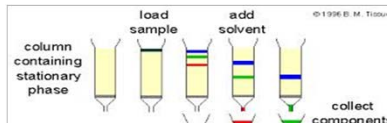
Protein Load Co-efficient Plot
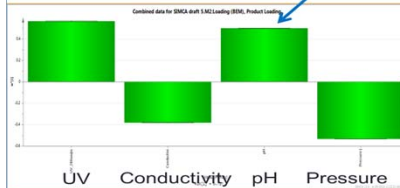
Elution Load Co-efficient Plot

BATCH LEVEL MODEL (BLM)

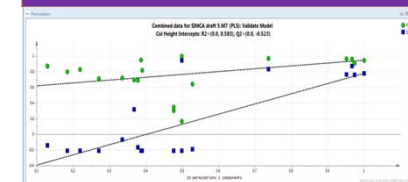

R2Y: 0.94
Q2: 0.725

CONCLUSION
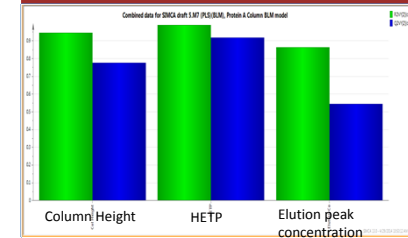
➢ Column height, HETP and elution peak significantly contribute to the performance of batches

VALIDATION

POTENTIAL APPLICATIONS

➢ Loading and elution behavior of current batches could be predicted.
➢ The resin shelf-life & online qualification of column could be performed.
➢ Process characteristics and resin comparison could be compared and assessed.
➢ DOE studies could be complemented to optimize the column dimensions during Tech transfer and scale-up.

ACKNOWLEDGEMENT