2014

Development of a High-Throughput Self-Interaction Chromatography System

David Johnson
University of Alabama at Birmingham

DEVELOPMENT OF A HIGH-THROUGHPUT SELF-INTERACTION CHROMATOGRAPHY SYSTEM

by

DAVID H. JOHNSON

LAWRENCE J. DELUCAS, COMMITTEE CHAIR
MARTHA W. BIDEZ
CHRISTIE G. BROUILLETTE
RICHARD A. GRAY
YUHUA SONG
W. WILLIAM WILSON

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2014

DEVELOPMENT OF A HIGH-THROUGHPUT SELF-INTERACTION CHROMATOGRAPHY SYSTEM

DAVID H. JOHNSON

BIOMEDICAL ENGINEERING

ABSTRACT

The first recombinant human protein drug, insulin expressed in e.coli cells, was approved by the FDA in 1982. Since then, protein therapeutics have become the fastest growing segment of the pharmaceutical industry and include immunoglobulin-g (IgG) directed cancer and immune disorder treatments. A major difficulty to bring protein drugs to market is the requirement that they be concentrated up to 150 mg/ml without aggregation for efficacy in a small injection volume. One way to improve protein drug solubility is to include additives that reduce protein-protein attraction and increase protein-protein repulsion thereby preventing protein molecules from coming together to form aggregates. However, hundreds of individual additives are approved by the FDA for injection and just ten additives at four possible concentration levels provides over a million (4^{10}) possible formulations. To address this formulation search problem, automated hardware and screening techniques are applied to evaluate the effect of additives on protein-protein interactions. These interactions are quantified by the second virial coefficient (B value), a thermodynamic parameter that is the sum of forces between two protein molecules at all orientations and distances in a solution. B values are measurements made by self-interaction chromatography. Contributions to the

formulation search problem include hardware, software and screen methodology improvements. The hardware consists of 1) robotic formulation delivery and system equilibration, 2) a reduced-cost flow cell design with protein detection utilizing UV LEDs and 3) a multi-column system for parallel experimentation. The software includes two parts 1) automation of the self-interaction chromatography experiment and 2) a neural network model of additive influence on protein-protein interactions. The hardware and software components are utilized in a tiered additive screen including individual additive evaluation (initial screen), a complex formulation evaluation (incomplete-factorial) and training of a neural network to model the additive influence on protein-protein interactions. The neural-network model is then used to predict the B value of additive combinations not previously measured and the prediction capability of the model is evaluated. The system was evaluated using an IgG drug candidate protein from Minerva Biotechnologies and predictions from the neural network produced a 100 fold increase in solubility.

Keywords: pharmaceutical formulations, protein solubility, physical stability, high-throughput screening, neural network, design of experiments

TABLE OF CONTENTS

Page

LIST OF TABLES

Table *Page*

PROTEIN SOLUBLIZATION: A NOVEL APPROACH

LIST OF FIGURES

Figure *Page*

INTRODUCTION

HIGH-THROUGHPUT SELF-INTERACTION CHROMATOGRAPHY: APPLICATIONS IN PROTEIN FORMULATION PREDICTION

HAZARD ANALYSIS AND RISK ASSESSMENT IN THE DEVELOPMENT OF BIOMEDICAL DRUG FORMULATION EQUIPMENT

PROTEIN SOLUBLIZATION: A NOVEL APPROACH

LIST OF ABBREVIATIONS

SIC: self-interaction chromatography

HSC: high-throughput self-interaction chromatography

DSC: differential scanning calorimetry

SLS: static light scattering

DLS: dynamic light scattering

FDA: Food and Drug Administration

ANN: artificial neural network

IgG: Immunoglobulin G

INTRODUCTION

Pharmaceutical Formulations and the Importance of Solubility

A new era of drug therapy began in 1982, with FDA approval of recombinant human insulin protein expressed in *Escherichia coli* cells[1]. The ability to deliver engineered proteins allows for replacement of missing protein and introduction of modified proteins to affect functional pathways. Therapeutic proteins are the fastest growing market segment of the pharmaceutical industry[2]. However, there are many challenges associated with protein drugs. High solubility is generally required for efficacy yet high-concentration protein formulations are typically prone to aggregation which reduces efficacy, shelf-life and has the potential to cause immunological response in patients[3–9]. Each potential protein drug requires the determination of a formulation to maintain its solubility and shelf-life. Additives can affect protein-protein interactions that influence aggregation and precipitation.

Pharmaceutical formulations may contain any additive provided it has been previously recognized as safe by the Food and Drug Administration (FDA). The FDA has approved over 200 unique additives that qualify as inactive and therefore "safe" ingredients[10]. Three to five unique additives (each at a different concentration) are often used in combination to prepare a useful formulation that provides optimum protein solubility and stability. If one considers the total number of possible combinations for

these 200 additives (i.e. the "space" of possible formulations, limited to just presence or absence of each additive) the mathematical result is 2^{200} formulations. This upper bound includes the unlikely formulation containing all 200 additives. However, if the choice is reduced to no more than 5 additives from the pool of 200 there are still over 2.5 billion options. Thus, it is simply not feasible to determine the effect of every additive combination on protein-protein interaction. The research presented here provides a system that reduces the number of additive combinations and concentrations that must be evaluated thereby providing a rapid method to determine the optimum formulation for a particular protein. In addition, this novel methodology and supporting automated technology reduces manual operations and the amount of protein required.

Design of biomedical instrumentation is inherently multi-disciplinary and includes hardware and software engineering design with consideration of system safety and efficiency as well as protein thermodynamics and formulation development. The project presented here uniquely addresses each component in the development of a high-throughput self-interaction chromatography system. Challenges to address safety, efficiency and potential failure of system components are addressed in "Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulations". Application of an artificial neural network to improve efficiency by minimizing of the number of physical experiments required is addressed in "Applications in Protein Formulation Prediction". A multi-tiered screening process identifies false positive results

due to protein denaturation and this full process is described in "Protein Solubilization: A Novel Approach". Each manuscript address a different aspect of the challenges presented in the design of an automated drug formulation system: 1) design for safety, efficiency and fault tolerance, 2) effects modeling and model validation and 3) quality control and system verification with increased solubility formulations for pharmaceutical proteins.

Protein Physical Stability

The physical stability of a protein molecule in solution depends on non-covalent intermolecular forces such as hydrophobic, electrostatic and Van der Waals forces all of which are influenced by additives contained in the solution[11,12]. There are many ways to assess physical stability -- both quantitatively and qualitatively. Qualitative methods include light obscuration, coulter counting, size exclusion chromatography (SEC) and dynamic light scattering (DLS)[13–15]. These methods determine the level of aggregation that exists in a solution. Quantitative methods, which can measure protein-protein interaction, include static light scattering (SLS)[16], analytical ultracentrifugation[17], osmotic pressure[18] and self-interaction chromatography[19].

Some form of particle visibility study is required for FDA approval of a protein pharmaceutical drug. The protein must remain in solution without forming visible

aggregates under stress. Stressors include high temperature and both high and low pH compared to storage conditions. This is the simplest method for determining protein aggregation. The protein is placed under a specific stress for a period of time and then visually inspected to assess aggregation[20]. Alternate methods exist to identify microscopic particles, including light obscuration, coulter counting and dynamic light scattering. Although each method can quantify a level of aggregation they do not quantify the underlying molecular forces involved.

The sum of forces between pairs of particles in a dilute solution is quantified by the thermodynamic parameter, second virial coefficient (B). McMillan-Meyer theory[21] applies the ideal gas law to canonical distributions of non-ideal gasses, taking into account interactions between particles. The theory is applied to solutions of large molecules by Zimm[22] and the formulation is given succinctly by Neal, et al[16]:

Eq 1. $P=RTc_p(1/M_W + Bc_p + \dots)$

In equation 1, P is the osmotic pressure (atm), R is the gas constant (L atm mol^{-1} K^{-1}), T is temperature (K), c_p is the protein concentration (g/L) and M_W is the molecular weight (g/mol). The second virial coefficient, B, (mol * ml / g^2 * 10^4) represents the sum of forces between 2-body interactions and the ellipses represent higher order, N-body, interactions. Without the B value and higher order components the formulation reduces to an ideal, non-interacting solution.

The second virial coefficient of a dilute protein solution can be measured directly via several methods, including static light scattering, sedimentation equilibrium, osmotic pressure and self-interaction chromatography.

Self-Interaction Chromatography

Self-interaction chromatography (SIC) is a technique to measure B values with reduced time and protein requirements compared to the methods described above. SIC is an affinity chromatography method in which the protein of interest is covalently bound to chromatography media and packed into a column. The formulation is flowed over the media and a bolus of protein is injected into the flow path. The amount of time required for the injected protein to elute from the column, compared to a non-interacting molecule, is directly related to the B value of the protein in the mobile phase solution. Lenhoff[19] describes the relation between B value and the ratio of interacting to non-interacting retention volumes as follows:

Eq. 2 B=NA/MW2(Vhs – k'/phi/rho)

Where k' is the retention factor – the ratio of interacting to non-interacting elution volumes centered about zero. NA and MW are Avogadro's number and molecular weight (mg/mL), respectively. Vhs is the hard sphere (excluded volume, mL)

contribution to B value. The ratio of volume to surface area of the media is phi and rho (mol/L) is the binding density of the protein on the media.

To prepare a self-interaction chromatography column requires less than 1 mg of protein per column. After protein is bound to the media, different formulations can be tested as the mobile phase in the system. The test of each additional formulation requires only 1 ug of additional protein injected into the system. Up to one-hundred formulations can be evaluated on a single column for an average protein consumption of 11 micrograms of protein *per B value determination* -- 10 micrograms prorated over the screen for the static phase and 1 microgram for the mobile phase .

A chromatography based system provides additional advantages besides its reduced protein consumption. As a standard technique for protein purification, chromatography systems are nearly ubiquitous in molecular biology labs. The operation of a liquid chromatography system does not require specialized expertise as is required for static light scattering.

HSC System

The HSC System is a novel, high-throughput, automated platform for B value measurement of a set of formulations. The system consists of four primary components: formulation delivery, injection system, data acquisition and control software.

Formulation delivery is provided by a syringe pump (formulation pump) and valve (formulation valve) which directs the pump to a reservoir to withdraw formulations into the syringes and then directs the pump to the rest of the system to deliver formulations.

The injection system consists of a second syringe pump (injection pump) which contains the protein of interest at a concentration of approximately 1 mg/ml and a valve (injection valve) that allows flow from either the formulation pump or the injection pump. During injection the formulation flow is halted, the injection valve is switched from formulation flow to injection flow and 1 ul of the protein is injected into the system. After injection, the injection valve switches back to formulation flow and the formulation pump continues to flow the remaining formulation (and protein injection) over the columns and through a flow cell.

Data acquisition operates at the flow cell downstream of the column. The flow cell contains a UV light source and UV diode detector. Similar to all HPLC systems the UV source and detector are separated from the flow by a quartz window and the protein concentration between the two windows follows Beer-Lambert law. The first prototype uses flow cells and a light source from Ocean Optics with the light delivered by fiber optic cable. The second prototype uses a custom flow cell design with direct UV LED illumination and flow through 0.02” i.d. peek tubing.

Custom control software handles pump and valve control for formulation delivery and protein injection. Data management is handled by the control software via a Measurement Computing (USB-1608FS) analog to digital converter. Incoming data from the system includes the UV transmission voltage from the UV diode detector and the pressure of each column as measured by four DJ Instruments DF2 pressure transducers.

Pressure sensing and detection during a screen is one of three major improvements driven by a failure mode effects and analysis and reported in the manuscript, "Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulations". The pressure sensor is necessary to detect aggregation of protein on the columns (a potential high risk failure because it can ruin a column containing milligram quantities of protein). This potential hazard also necessitates the ordering of additives such that lower concentrations of individual additives are evaluated before complex formulations. Another common failure mode in chromatography systems – the introduction of air bubbles into the system – is addressed by including bypass valves. The addition of bypass valves allows the system to be flushed at higher flow rates to move air bubbles through the system without producing pressure in the columns. Although this configuration is not novel to HPLC systems the automated control of the HSC system means that an unattended failure due to air would be damaging to multiple experiments. Automated clearing of air bubbles as part of the protocol developed to change

formulations between experiments is one of the ways identified to mitigate the potential hazard.

The ability to rapidly measure B value of a protein in different formulations allows for a more comprehensive screening process with hundreds of physical B value measurements. In order to optimize use of the instrument a multi-tiered screen is used to systematically search for additives and combinations of additives which increase B value formulation. The search space of possible additives is first restricted by evaluating a formulation with a single additive. Measurement of B value with single additive formulations allows for the ordering of many additives due to only needing a single machine run per additive. However, this does not reveal anything about how additive concentration affects B value nor anything about the effect of additive combinations. Therefore, the next screen tier uses more complex formulations with two to three additives in combination and at varying concentrations -- chosen from the most promising additives of the initial screen. The experiment design for this screen is based on an orthogonal array to ensure a distributed sampling throughout the possible space of formulations. After two tiers of physical B value measurements the effect of additives on B value is modeled with an artificial neural network. The model is used to generate (predict) all combinations of additives and concentrations in the space of formulations that was originally sampled with the orthogonal array. The neural network model is evaluated by physically measuring the formulation B values using the HSC system.

Details of the artificial neural network model are given in “Applications in Protein Formulation Prediction”. The multi-tiered screen process is described in more detail in “Protein Solubilization: A Novel Approach”. The remaining methods detail the individual techniques used with the instrument to acquire B value measurements.

HSC Methods

Self-interaction chromatography (SIC) is an affinity chromatography technique used to determine B value (integral sum of all 2-body interactions) of a protein in a test formulation. The protein of interest is bound to chromatographic media as the stationary phase. The test formulation is flowed through the system as the mobile phase and a bolus of the protein is injected into this mobile phase. The elution volume of the protein is indicative of the protein self-interaction (or cross-interaction for different immobilized and injected molecules). The working equation [Tessier, Lenhoff] for the B value calculation is:

Eq 3. $B = \frac{N_A}{MW^2}\left(B_{HS} - \frac{k'}{\phi\rho}\right)$

Where N_A is Avogadro’s number and MW is the molecular weight of the protein (g/mol). Inside the parenthesis, B_{HS} is the hard sphere contribution of the excluded volume (based on a globular model of protein volume as a function of molecular weight) and ρ is the amount of protein bound to the column in molecules per unit area

(molecules/cm^2). The phase ratio, φ, is the ratio of surface area to volume of the chromatographic media (cm^2/ml) (dependent on the media and protein volume) and is calculated according to the experimental characterization of the phase ratio of chromatographic media by DePhillips and Lenhoff[23]. All of these variables are constant for a given column with bound protein. The retention factor, k', is a standard chromatographic measure, which is a function of the retention volume of the protein and the retention volume of a hypothetical non-interacting molecule of the same size. Note that if the retention factor is negative, then the overall B value is positive. The retention factor is calculated by the formula:

Eq 4. $k' = \frac{V_p - V_0}{V_0} = \frac{V_p}{V_0} - 1$

In this equation, V_p is the retention volume of the protein and V_0 is the retention volume of an equivalent, but non-interacting molecule. Obviously, it is not possible to construct an exactly equivalent non-interacting marker, so there are two alternative ways to calculate the non-interacting volume: 1) an acetone marker (2% v/v) or 2) the protein injected over a "dead" column which has no protein bound (and has been capped to prevent protein binding during the experiment).

In review, the SIC experiment requires the following steps:

1) Bind the protein to chromatographic media
2) Cap chromatographic media

3) Pack media into column and determine the concentration of bound protein by BCA assay
4) Attach the column to the HSC system
5) Flush the system with the formulation of interest
6) Inject a bolus of the protein into the mobile phase
7) Measure the retention volume of the protein at fixed flow rate (8 ul/min)
8) Determine the non-interacting volume
9) Calculate B value

The first four steps are part of column preparation and the last four steps are part of the HSC operation. More detail regarding both parts will be covered in the next two sections.

Column Preparation

Four media types designed for protein binding are available from Tosoh Haas, each with different binding chemistry. The media names are given by the active binding group: formyl, tresyl, carboxy and amino. During the initial evaluation of a protein, different binding chemistries are tested in small batch (20 ug of media, 20 ul protein) to identify the chemistry with the highest binding affinity for the specific protein. The

media with the highest binding affinity is chosen for the stationary phase in the HSC system. The available resins are given in the following table:

Table 1. Available chromatographic media

Media	Binding Group	pH	Buffer	Reagents	pH	Capping
Tresyl	-NH2,-SH	8	PBS	None	8	0.5M Ethanolamine
Formyl	-NH2	8	PBS	NaCNBH4	8	0.5M Ethanolamine
Amino	-COOH,-COH	6	MES	EDC/NHS	6	0.5M 3HP
Carboxy	-NH2	6	MES	EDC/NHS	6	0.5M Ethanolamine

For binding, it is important to use a formulation that does not contain the active binding groups in the additives. We use a formulation as close as possible to the storage formulation (received formulation) excluding any additives (or buffer systems) that contain the binding groups. The binding concentration for each media is determined via Pierce BCA assay using bovine serum albumin (BSA) for the standard curve. The media that binds at the highest concentration is used for all subsequent large scale media preparations. The preferred binding concentration is at least 5 mg/ml.

For SIC columns, the protein is bound to the media and capped in bulk. For the "dead", non-interacting, column a separate batch of media is subjected only to the capping process. Activation reagents for the different media (if necessary) are given in Table 1. After media is capped (with or without the protein binding step) it is packed into tubing 18 cm long, 0.5 mm i.d. Two additional cm of column are packed and cut from the end. These two clippings (1cm each) are subjected to a BCA assay to determine the

amount of protein bound in the packed column. The beads are agitated in the BCA reagents to free the beads from the column clippings. The concentration of bound protein (in mg/ml) is then used to determine ρ (molecules / area) for the B value calculation.

HSC Operation

The HSC system has four separate channels each with a UV280 detector and injection syringe. In addition to separate channels there is also a formulation reservoir connected to a robotic platform that provides automated system washing (via column bypass) and replacement of formulations. The layout of the four channels on the machine is given in Figure 1.

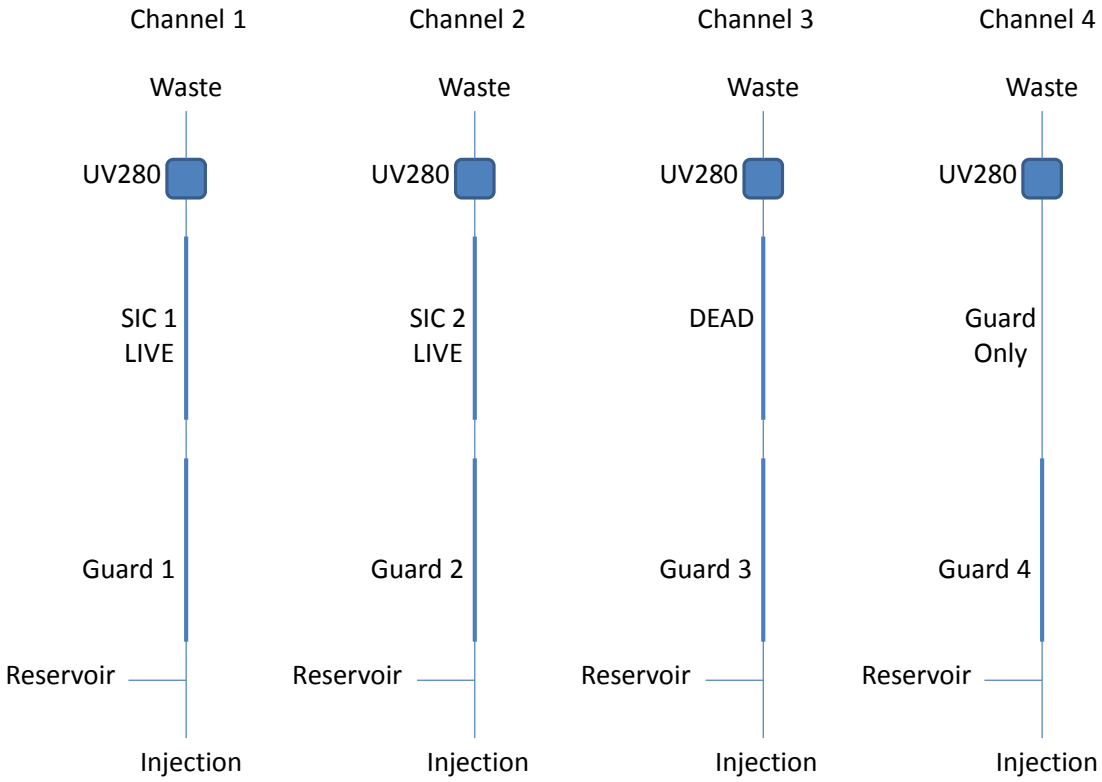

Figure 1. HSC Channel Layout

Note that guard columns, not previously discussed, are present immediately after the injection port. The guard column, packed with Sephadex-G15, is used to ensure that the protein traveling through the SIC column is equilibrated in the mobile phase formulation and separated from injection formulation components (including the acetone marker).

An injection formulation containing the protein and 2% acetone (v/v) gives an ideal chromatogram consisting of two separate gaussian peaks. The first peak is the

protein molecule and the second is the acetone marker. The MW cutoff of the Sephadex G15 media is 1.5 kDa, therefore any protein above that molecular weight is expected to be separable from the injection formulation. A minimal formulation is identified during the pre-screen that produces two gaussian peaks over guard and dead columns. This minimal formulation is used as a baseline formulation throughout the screening process to ensure the system gives consistent retention times.

Figure 1 represents the channel layout during measurement of retention volumes used to calculate B values. Prior to the layout pictured in Figure 1, the volume required to elute both protein and acetone through the system without a SIC column is determined experimentally by injection of a bolus of protein into the reference formulation with only guard columns attached in all four channels (as in channel 4). This gives the "system volume" for the protein and acetone which is specific to each channel and guard column. The system volume per channel is subtracted from each retention volume measurement to determine protein and acetone retention volumes through the SIC columns. Subtraction of the system volume, V_s, from each retention volume results in an updated equation for retention factor:

Eq 5. $$k' = \frac{(V_p - V_s) - (V_0 - V_s)}{(V_0 - V_s)} = \frac{(V_p - V_s)}{(V_0 - V_s)} - 1$$

Expected retention times for guard only elutions are 1000 s for protein and 1700 s for acetone. If a peak is non-gaussian or the retention times are outside a +/- 200 s range

then the injection is considered failed. Figure 2 gives an ideal chromatogram for guard only elutions.

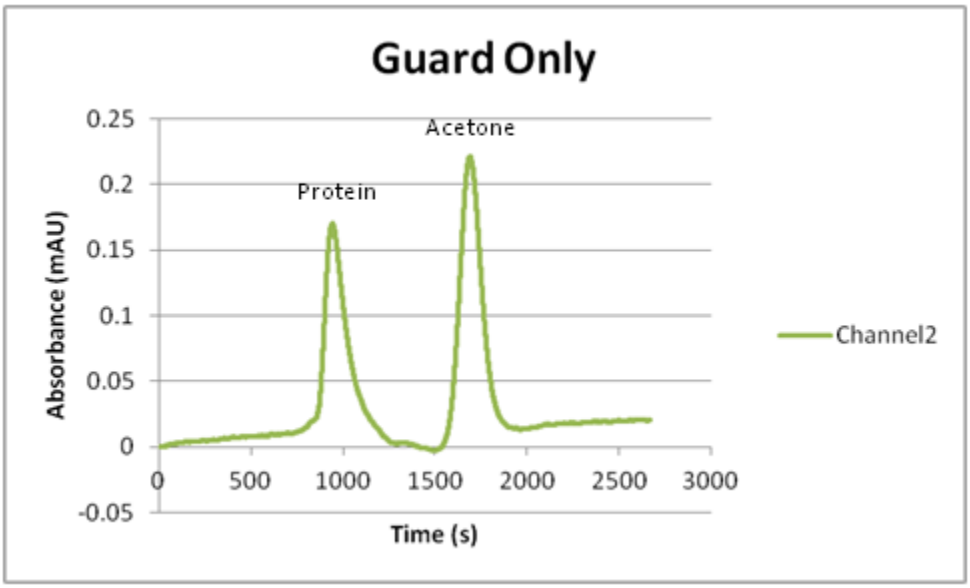

Figure 2. Ideal chromatogram for guard only elutions.

After measurement of system volume, the system is set up according to Figure 1 with live columns (protein bound) connected to channels 1 and 2, a dead column (no protein) connected to channel 3. Channel 4 remains guard-only to measure system volume excluding the self-interaction chromatography column.. In the case of a failure the chromatogram of channel 4 gives an indication of whether interactions between the protein and guard media contribute to the failure. However, after identification of a minimal formulation, we generally do not see failures of this type. Channels 1 – 3 are used to calculate B values. Figure 3 represents an ideal chromatogram for live column

elutions. Note that this chromatogram is almost identical to the guard-only chromatogram except that the peaks elute ~400 seconds later.

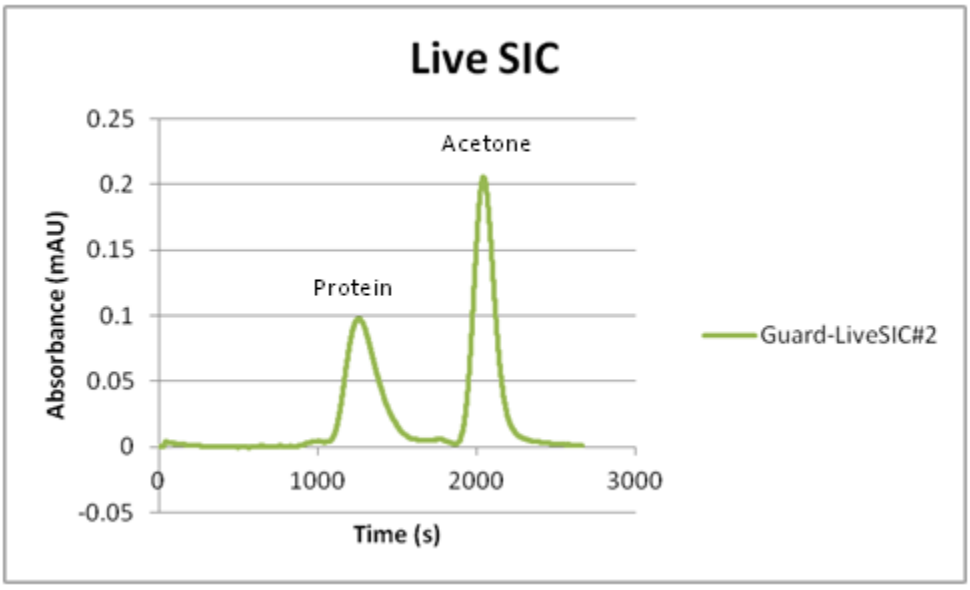

Figure 3. Ideal chromatogram for live column elutions.

*B value calculatio*n. Retention times through the SIC column are calculated for each channel by subtracting guard only retention times from the experimental retention time. These retention times through the SIC column are used to calculate retention factor. Note that, at constant flow provided by the formulation syringes, a ratio of retention times is equivalent to a ratio of retention volumes. Therefore, retention times are used in place of the physical retention volumes when calculating the retention factor, k'.

The two methods to calculate B value differ in which retention volume is used for the "non-interacting" retention time for calculation of the retention factor. Either the acetone retention volume or the dead column retention volume can be used. The method reported in the literature uses acetone as a non-interacting marker. This is how we calculate and report B values. However, we suspect that the dead column may be a better representation of the non-interacting retention volume in the case where protein-media interactions are significant. This has been observed when an additive causes denaturation of the protein in aggressive solubility screens. We are in the process of identifying formulations which give significantly different B value measurements between the acetone volume method and dead volume method. Evaluation of these formulations by an independent B value measurement, such as static light scattering will be able to identify which method is more accurate. Until a definitive conclusion is reached, the policy of Soluble Therapeutics is to calculate and report B values using an acetone marker as non-interacting, but to also collect and calculate B values using the dead column as non-interacting. Table 2 shows typical parameters and retention times when calculating B values for IgG proteins.

Table 2. Typical B value parameters for IgG

Parameter	**Value**	**Units**	**Description**
Binding Concentration	5	mg/ml	Minimum
Phase Ratio (phi)	5.79	m^2/ml	Media Area/Volume
Molecules / Area (rho)	2.94E+15	molecules/m^2	IgG at 5 mg/ml

Retention Factor (k')	0.012	ratio	non-interacting
Bhs	7.16E-19	ml/molecule	for IgG
Protein RT Guard-SIC	1250	seconds	Typical
Protein RT Guard	950	seconds	Typical
Protein RT SIC Only	300	Seconds	Typical
Acetone RT Guard-SIC	2000	Seconds	Typical
Acetone RT Guard	1700	Seconds	Typical
Acetone RT SIC Only	300	seconds	Typical

Validation – During Screen

B value deviations in the minimal formulation greater than 1 B unit (mol*ml / g^2) or retention time deviation in the minimal formulation greater than 10% result in re-evaluation of the formulation immediately after the failed minimal formulation and column replacement if the deviation persists through more than one minimal formulation.

Validation – Post Screen

After the initial screen the nine formulations with the most positive B values are evaluated by dynamic light scattering (DLS) and differential scanning calorimetry (DSC) to confirm that the formulation does not give a false positive B value, actually producing aggregates or denaturing the protein. DLS results that indicate a majority of scattering intensity is due to higher order aggregates than observed in the minimal and storage

formulations are considered failures. DSC results that indicate a denatured protein (no signal) or an onset of unfolding less than 37 ^{o}C are considered to failures. Any formulation that fails this step does not proceed to the discovery screen (incomplete factorial) and is replaced by the next most positive B value measured (which are also validated by DLS and DSC).

The high-throughput and automated HSC system provides a platform to screen a large number of formulations. Selection of the additives to screen is determined in three phases: initial screen, incomplete factorial screen and neural-network screen. The initial screen consists of individual additives to identify those single additives which most reduce protein-protein interactions. Additives which give the nine most positive B values are chosen for the incomplete factorial screen. To test all additive combinations at all concentrations would require millions of formulation conditions. Therefore the incomplete factorial screen takes an orthogonal subset of the formulations – orthogonal so that each additional formulation tested provides additional information about the effect of the individual additives and how they interact with other additives. Results of the incomplete factorial screen are used to train an artificial neural network for the third phase.

Two key aspects of the incomplete factorial design contribute to success of neural network training. The first is the basis on an orthogonal array. The orthogonal array provides a balanced representation of each additive with each other additive within the

screen. This serves to maximize the information available to the neural network for each additive. Another factor in the success of the incomplete factorial is pre-selection of additives based on the initial screen. Additives which have been pre-selected for positive B value (and screened against denaturation using DSC) are less likely to cause chromatographic problems due to aggregation.

The artificial neural network is a model of how different additives affect B value. Training of the ANN is performed using back-propagation. Initial weighting of the neural network is randomized and a B value is calculated for each formulation using the random weighting. The error is calculated between the neural network output and the measured B value. This error is used to adjust the weights based on their relative contribution to the error. This process is repeated to reduce error for a subset of the formulations with measured B value. Overfitting of the network to the error in the training set is avoided by monitoring the training process with a validation set. The ANN is used to calculate B values for a set of formulations not previously used to modify weights. As long as the training process produces a decrease in error of the validation set the training continues. When weight modification based on the training set data produces an increase in error of the validation set the training is halted and the previous weight set is identified as optimal.

Three manuscripts are presented which track development of the HSC System through concept, design and implementation. The first publication, “Applications in

Protein Formulation Prediction" focuses on the neural network aspect of the system. That is the ability of a neural network, when trained on measured B values of an incomplete factorial screen, to predict B values of formulations previously not measured by self-interaction chromatography. The second publication, "Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulations" covers the design of the instrument and analysis of potential hazards, both operational and physical in creating the HSC system. The final publication, "Protein Solubilization: A Novel Approach", demonstrates the full screening process using the HSC instrument on a pharmaceutical drug target. The screen process is able to improve solubility of the target protein one-hundred fold.

HIGH-THROUGHPUT SELF-INTERACTION CHROMATOGRAPHY: APPLICATIONS IN PROTEIN FORMULATION PREDICTION

by

DAVID H. JOHNSON, ARUN PARUPUDI, W. WILLIAM WILSON, LAWRENCE J. DELUCAS

Pharmaceutical Research **26,** 296–305 (2009)

Format adapted for dissertation

Abstract

Purpose. Demonstrate the ability of an artificial neural network, trained on a formulation screen of measured second virial coefficients, B, to predict protein-protein interactions for untested formulation conditions.

Materials and Methods. Protein-protein interactions, quantified by the second virial coefficient, B, are measured by self-interaction chromatography (SIC). The B value of lysozyme is measured using an incomplete factorial of 81 formulation conditions. The influence of screen parameters (pH, salt, additives) on B value is modeled by training an artificial neural network on B value measurements. After training, the neural network is asked to predict the B value for the complete factorial of parameters screened (12,636 conditions). Twenty of these predicted values (distributed throughout the range of predictions) were experimentally measured for comparison.

Results. The neural network was able to predict lysozyme B values with a significance of $p<0.0001$ and RMSE of 2.6 mol ml / g^2.

Conclusions. The results indicate that an artificial neural network trained on measured B values for a small set of formulation conditions can accurately predict B values for untested formulation conditions. As a measure of protein-protein interactions correlated with solubility, B value predictions based on a small screen may enable rapid determination of high solubility formulations.

Introduction

A protein's interaction with itself and with other proteins affects important characteristics such as its solubility[1], aggregation[2] and ability to crystallize[3]. Measurement of the second virial coefficient, B[4], provides one method to quantify protein interactions at the molecular level. B is a measure of the entirety of two body (protein-protein) interactions that includes contributions from excluded volume, electrostatic factors (attractive and repulsive) and hydrophobic interactions. In terms of McMillan-Meyer solution theory[5], B is related to a potential of mean force, which describes all of the interaction forces between two protein molecules in dilute protein solution. Positive B values correspond to net repulsive forces between proteins and are correlated with increased protein solubility in solution[1, 6] whereas values in the negative range correspond to the net attractive forces required for protein crystallization[3]. Identified as one indicator of the physical stability of proteins in solution[7], the second virial coefficient depends on a variety of solution formulation parameters including temperature, pH and the type and concentration of salts and excipients (additives).

The initial evaluation of a protein's function in human pathology is often facilitated by study of the protein's structure by means of x-ray diffraction. The second virial coefficient can provide functional insight at every step in the drug discovery process. George and Wilson have shown[3] that proteins generally will crystallize when their B values are in a "crystallization slot" ranging from approximately -0.2 to -8 (x 10^{-4} mol ml/g^2). This B value range represents slightly to moderately attractive forces between proteins, a condition that appears to be important for nucleation and subsequent crystal formation. The George and Wilson crystallization slot has been confirmed by

several research groups on a variety of different proteins[2, 8-10]. Crystallization conditions which fall within the crystallization slot do not necessarily produce protein crystals. However, crystallization conditions which fall outside the crystallization slot have protein interaction dynamics in which proteins either come together too quickly to form unordered precipitation or do not come together at all. Knowledge of a protein's B value in a crystallization condition could therefore be used to reduce the number of false leads pursued by crystallographers by the exclusion of conditions that result in protein B values that fall outside the crystallization slot.

The determination of solution conditions yielding diffraction quality crystals as well as low solubility and/or unwanted nonspecific aggregation of proteins expressed in prokaryotic and eukaryotic systems represent major bottlenecks in high-throughput protein structure determination[11, 12]. Although there have been advances in the ability to recover bioactive protein from the inclusion bodies of various expression systems[13] these techniques require customization to the protein of interest, a requirement that is not conducive to high throughput methods. The mathematical relationship between the B value and solubility, derived by Haas et al[1], indicates a marked increase in solubility with increasing B value. This relationship has been validated experimentally for a variety of proteins[1, 6, and 8]. Thus, a second application of the second virial coefficient involves its use as a diagnostic for protein solubility.

In addition to the role of solubility in protein expression and structural studies, various human pathologies are directly linked to solubility and protein aggregation. Alzheimer's disease, Parkinson's disease[14] and cataracts[15] are some examples of disease pathologies caused by abnormal protein aggregation. Measurement of protein-protein

interactions quantified by the second virial coefficient has been used to determine aggregation likelihood[16]. This implies that the second virial coefficient could be used to directly assess the effect of therapeutic agents (both protein and small molecule) on the protein-protein interactions that result in specific disease pathologies.

Protein solubility and stability are as important in the evaluation of therapeutic proteins as they are to the study of proteins involved in disease pathology. Food and Drug (FDA) evaluation of a drug candidate includes two primary criteria: solubility and membrane permeability[17]. Not only is solubility required for membrane permeability, but orally active (i.e. drugs taken by mouth as opposed to injection) candidates require an even higher level of solubility. In a recent overview of pharmaceutical drug screening techniques[18] three methods of solubility screening were identified: UV absorption, nephelometry and flow cytometry. These methods, developed for analysis of small molecules, are used to calculate current or potential solubility of a specific drug formulation and can be performed in a high throughput manner. However, they do not directly quantify the protein-protein interactions that influence solubility and aggregation of protein therapeutic molecules. A high throughput screen of second virial coefficients would enable rapid quantification of protein-protein interactions for a large number of formulation conditions.

Measurement of the second virial coefficient is traditionally performed using static light scattering (SLS)[3]. SLS is an application of the same light scattering phenomenon used in nephelometry for determining solubility. However, with SLS the scattered light is measured in increments of known protein concentration which enables calculation of the underlying second virial coefficient. This traditional method of

determining B consumes a significant amount of protein and time (multiple light scattering readings are necessary to calculate one B value) and it requires careful attention to solution clarity.

In contrast, a second method for determining the second virial coefficient known as self-interaction chromatography (SIC) provides advantages to each of the constraints of the SLS method referenced by Tessier et al. SIC initially requires chemical coupling of protein to a solid support followed by careful packing of the support in a small chromatography column (70 mm long, 0.4mm ID). Once prepared, however, the column is very stable and can be repeatedly used to measure B values about every 8 minutes, making it more applicable to high-throughput techniques. Each measurement consists of flowing a mobile microgram injection of the protein across the immobilized protein particles using a high-precision HPLC. The retention time of the mobile protein is directly related to its interaction with immobilized column protein[19], thereby providing a direct measurement of how two proteins (bound and injected) interact with one another. Formulas relating the chromatographic retention time to B values can be found in Tessier's work[20]. This technique has been successfully used with low throughput screens (16 conditions) to measure the interactive effects of two formulation parameters on B[21].

In this study we measure the B value for 81 solution formulations in a high throughput screen using self-interaction chromatography. The screen measures the pair-wise effects of nine different additives on the self-interaction of the protein hen egg-white lysozyme. The well known incomplete factorial experimental design technique, which was applied to crystallization screening by Carter[22], is used to ensure wide coverage of the search space with a reduced number of test conditions. The incomplete factorial

design is accomplished by mapping the parameters of interest (pH, salts, additives, concentrations) onto an orthogonal array[23,24]. Mapping parameters to an orthogonal array allows equal representation of parameter levels throughout the search space while reducing the over 12,636 possible parameter combinations down to a reasonable screen size of 81 conditions. The B values are measured to quantify the degree of lysozyme self interaction in each formulation.

The results of the screen are first analyzed by manually examining the linear and quadratic trends of each formulation parameter on B value. The parameters with the most statistically significant effect on protein-protein interaction (B value) of lysozyme are identified within the screen. These parameters with strong influence on protein interactions (such as NaCl) are shown to have an effect on B value regardless of the presence of other additives in varying formulations. This allows for the rapid identification of additives which could be used to modify protein-protein interactions.

While a manual examination of parameter effects can identify the strong correlations of single parameters this initial analysis does not examine the effect of parameter interactions. To analyze the effect of additive combinations on protein-protein interaction we model the results of the B value screen using an artificial neural network (ANN). Artificial neural networks have utility when the effect of specific combinations of a large number of variables/parameters as well as each variable's level (i.e. concentration of various chemicals) must be analyzed to determine the optimal combination to yield a desired outcome. The large number of potential additive combinations and their possible levels defines a search space that precludes manual inspection of the data as a reasonable method for finding the optimum parameters and

parameter concentrations. Artificial neural networks are able to utilize an incomplete factorial subset of parameter combinations to determine correlations between discrete variables combinations and their respective levels. Neural network modeling has been used to predict novel crystallization conditions[25] and to confirm theoretical calculations of B for very small molecules[26].

An ANN is essentially a set of non-linear weighted functions which map input variables (screen parameters) into output variables (B value)[27]. The weights are initialized to random values which results in a random mapping of the screen parameters onto B values. The subsequent process to determine optimal weights and training is performed by iteratively updating the weights to reduce error between the ANN output and observed values (B screen). For each iteration, the ANN attempts to produce B values closer to the observed B values for the given input parameters. After the training process is complete, the neural network model is used to produce B value predictions for all possible formulations of one or two additives.

ANN B value predictions for twenty different formulation conditions were experimentally validated via SIC B values of lysozyme dissolved in each condition. The chosen conditions included ten from the most positive and negative B value predictions combined with ten spread throughout the range of predicted B values. The results demonstrate that an artificial neural network trained using an incomplete factorial subset of a total additive screen can accurately predict the second virial coefficient of previously untested formulations.

Finally, the ANN model is compared with a more traditional generalized linear model (GLM). Identical parameters used as inputs for the artificial neural network are included for consideration by the GLM. In the stepwise procedure the GLM uses an iterative process to determine which parameters significantly influence the second virial coefficient. For GLM analysis, the gradual process of ANN weight determination during each iteration is replaced by linear regression to calculate optimal linear model coefficients. The significance of each parameter is considered during each iteration, with new parameters added or removed based on a predetermined alpha value threshold. Once significant parameters are identified by this stepwise process, linear model coefficients are calculated by linear regression. The GLM, like the trained ANN, can be used to predict the B values of untested formulation conditions. Comparison of the GLM predictions to the ANN predictions indicates that, for this application, the ANN produces a more accurate and robust model than the GLM.

Materials and Methods

Screen Conditions

Hen egg-white lysozyme was purchased from Calbiochem. The chromatography particles, Toyopearl AF-Formyl-650M, were purchased from Tosoh Bioscience. Buffer formulation chemicals include glycerol, glycine, glutamic acid, mannitol, sodium citrate, sodium acetate and acetic acid; all purchased from Fisher Scientific. Additional formulation chemicals PEG4000, MPD and trehalose were purchased from Sigma-Aldrich under the Fluka brand name. Sigma-Aldrich was also the source for

chromatography bead capping agent, ethanolamine, as well as the formulation chemicals Na_2SO_4, Na HEPES, HEPES acid and citric acid. The final two formulation chemicals, succinic acid and arginine, were purchased from Acros Organics.

Each of the 81 solution formulations contain buffer, salt and one or two co-solvents listed in Table I. The identity of each formulation was determined by mapping these parameters onto an orthogonal array design as described by Sloane, et. al.(Sloane). This mapping produces formulation targets in which each pair of variables are equally represented throughout the screen (thereby producing a balanced screen with respect to the influence of individual parameters).

Table 1. List of additives, salts and buffers utilized in the formulation screen.

Buffers	Salts	Additives		
Acetate (pKa 4.7)	NaCl	Arginine	Sucrose	MPD
Succinate (pKa 5.6)	NaCitrate	Glutamic Acid	Mannitol	PEG4000
MES (pKa 6.1)	Na_2SO_4	Glycine	Trehalose	Glycerol
HEPES (pKa 7.5)				

The water source for formulations was pre-filtered at 18MΩ by a Millipore MilliQ system with trace sodium azide added to retard bacteria growth. Sodium and acid forms of 0.1 M buffers are mixed at their pKa in the presence of co-solvents (except in the case of the succinic buffer which was adjusted to pH with NaOH). The pH of each solution was confirmed via a Corning 430 pH meter with the final pH adjusted solutions filtered (0.22µm (Fisher Scientific) syringe filter) and stored at room temperature.

Protein Immobilization

Lysozyme (LYZ) was immobilized to AF-Formyl-650M beads as described by Valente et al[28] with only slight modification. One ml of 1M K_2HPO_4 at pH 7.0 was added to 350µl of AF-Formyl-650M beads followed by centrifugation (bench-top, 30 seconds 7k rpm). The wash was performed an additional two times to remove excess packing buffer. LYZ (5mg) was dissolved in the phosphate buffer and incubated with the beads. Fifteen mg of sodium cyanoborohydride was added to the bead mixture to activate the binding chemistry and mixed via rotary mixer at room temperature for 90 minutes. A 5µl sample of the supernatant containing unbound LYZ was diluted with 45µl of 0.1M sodium acetate buffer pH 4.7 and assayed via a bicinchoninic acid (BCA) assay (Thermo Scientific). The beads were centrifuged and washed twice with phosphate buffer plus 5% (w/v) NaCl to remove any remaining LYZ. After binding and washing, unreacted formyl groups were capped by adding 1ml of 1M ethanolamine at pH 8.0 and 10mg sodium cyanoborohydride, followed by additional rotary mixing for 90 minutes. After this final step of immobilization the beads were washed twice with 1mL of the sodium acetate buffer.

Self-Interaction Chromatography

Immobilized beads were packed into a micro-column consisting of teflon FEP tubing (i.d. 0.03", o.d. 1/16") and blocked at one end by a stainless steel frit (Valco). Two 1.1cm lengths of the packed tubing (~5µl each) were cut from the packing end,

diluted with 45µl of sodium acetate buffer and assayed using the BCA assay (Pierce Biotechnology) to determine protein binding density on the column. The packing end of the column was then cut to 18cm length, sealed with an additional frit and stored with the 0.1M sodium acetate buffer pH4.7 at 4 °C when not in use. A second column, referred to as the dead column, was packed with beads that have been subjected to only the capping portion of the immobilization procedure. Acetone was used as a non-interacting void-volume marker and was dissolved in water at 3% (v/v) for injections. The protein injection solution consists of 5mg of lysozyme dissolved in 1ml of each of the four separate 0.1M buffers (Table I).

All chromatograms were generated using a High Performance Liquid Chromatography (Shimadzu) system consisting of two pumps, an autosampler for sample injection, column oven, 280nm UV detector and software for automatic retention-time calculation. Each screen formulation was run through the column at 60 µl/min and the auto-sampler was used to inject 1µl of the 5mg/ml LYZ solution in buffer identical to the formulation buffer applied to the column. Column temperature was maintained at 23 °C. Injections were performed in triplicate over the same column and B values measured for the entire 81-condition screen on two columns with final B values averaged for each column. Solutions with outlying (1.5*IQR) variance (N=9) between two columns were measured on a third column. If the B of two columns were within the average standard deviation between 2 columns (1.7 B units) the disagreeing measurement was excluded. Sample chromatograms shown in Figure 1 demonstrate the influence of NaCl on retention time measured at peak elution.

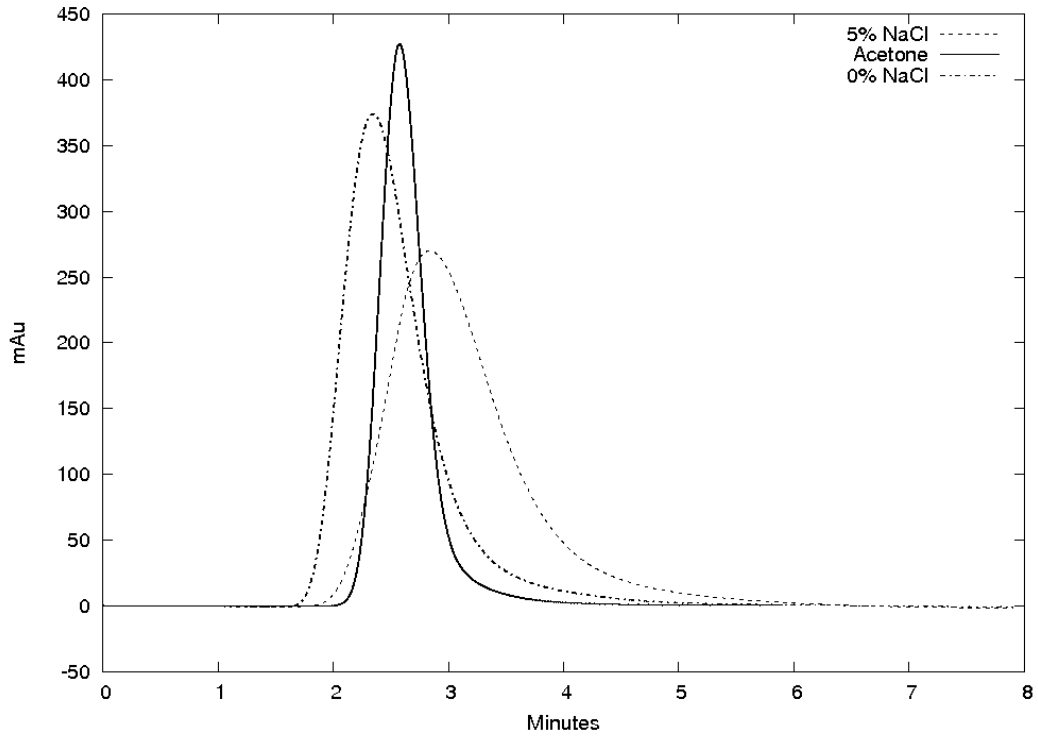

Figure 1. Retention times for lysozyme in 5% NaCl and 0% NaCl in 0.1M sodium acetate buffer demonstrates the affect of NaCl on lysozyme self-interaction. The retention time for 3% acetone in the same buffer with 5% NaCl provides a reference point for conversion of retention times to B values.

In the primary equation used to calculate B values, Eq. 1, N_A is Avagadro's number and MW is the molecular weight of the protein. The phase ratio, ϕ, is the ratio of the available surface area per unit of null volume and has been calculated for a variety of different chromatography particles[29].

$$B = \frac{N_A}{M^2}\left(B_H - \bar{s}\frac{k'}{\phi}\right) \quad (1)$$

The density of protein immobilized on the column is rho. The variable k' is the chromatographic retention factor calculated from the protein retention time (t_r) and acetone retention time (t_0) given by the equation:

$$k' = \frac{t_r - t_0}{t_0} \quad (2)$$

In this equation, Eq. 2, the acetone retention time (t_0) acts as a non-interacting marker to establish the relationship between non-interacting molecules with bound protein compared to interacting molecules with bound protein.

Static Light Scattering

The traditional static light scattering (SLS) experiment requires measurement of the scattered light intensity from a protein solution in excess of background as a function of protein concentration. We have modified the traditional SLS experiment in two important ways in order to minimize both time and protein required for a single B measurement[30]. The first modification is the incorporation of a low volume (~ 1 μL) scattering cell. The second modification is a configuration allowing the simultaneous measurement of scattering intensity and protein concentration. In short, this is accomplished by using a bifurcated fiber to deliver both the incident laser beam for scattering and the incident UV beam for absorption (protein concentration) measurements. The advantage of this configuration is that the simultaneous measurement of light scattering intensity and protein concentration allows the determination of the second virial coefficient from a single injection of protein sample into a flow system. Typically, 5 – 10 μl of protein solution at 1 – 2 mg/ml protein concentration were required for a single B measurement.

The intensity and concentration data were treated according to the SLS working equation[31]:

$$\frac{K\,c}{R_{90}} = \frac{1}{M} + 2B \qquad (3)$$

where K is an optical constant (cm^2 mol g^{-2}) given by $K = 4\,\pi^2 (dn/dc)^2 n_o^2 / (N_A \lambda^4)$, c is the protein concentration (g cm^{-3}), R_{90} is the Rayleigh factor (cm^{-1}) at angle 90 □, M is the molecular weight of the protein (g mol^{-1}), B is the second virial coefficient (mol ml g^{-2}), dn/dc is the refractive index increment (cm^3 g^{-1}), n_0 is the solvent refractive index, N_A is Avogadro's number (mol^{-1}), and □ is the wavelength vacuum. According to Eq. 3, a plot of Kc/R_{90} vs c (often called a single angle Zimm plot) linearizes the SLS data and B is determined from the limiting slope.

Artificial Neural Network (ANN)

Artificial neural network modeling was performed using the Java Object Oriented Neural Engine (JOONE)[32]. Figure 2a shows the overall network topology of the neural network used in this study including inputs, node configuration and B value output. Each node represents a nonlinear transformation of inputs and is grouped into one of two layers according to distance from the input parameters. Regardless of position in the topology, the output of each node is calculated by two steps shown in Figure 2b. First a weighted sum of inputs to the node is calculated, z. The hyperbolic tangent is taken of this weighted sum to calculate node output. Each node in layer 1 takes as input all formulation parameter while each node in layer 2 takes all outputs from layer 1 as input.

The final B value output is calculated as a simple weighted sum of layer 2 outputs without a nonlinear transformation. This permits the range of output values to match the range of screened B values rather than the (-1,1) range of the hyperbolic tangent function. Through calculation of each layer's outputs in sequence this architecture is able to estimate a B value for a given set of condition formulation parameters. The weights associated with each input node are the variables subject to training thereby creating a network function that most accurately represents B values over all given formulation parameters.

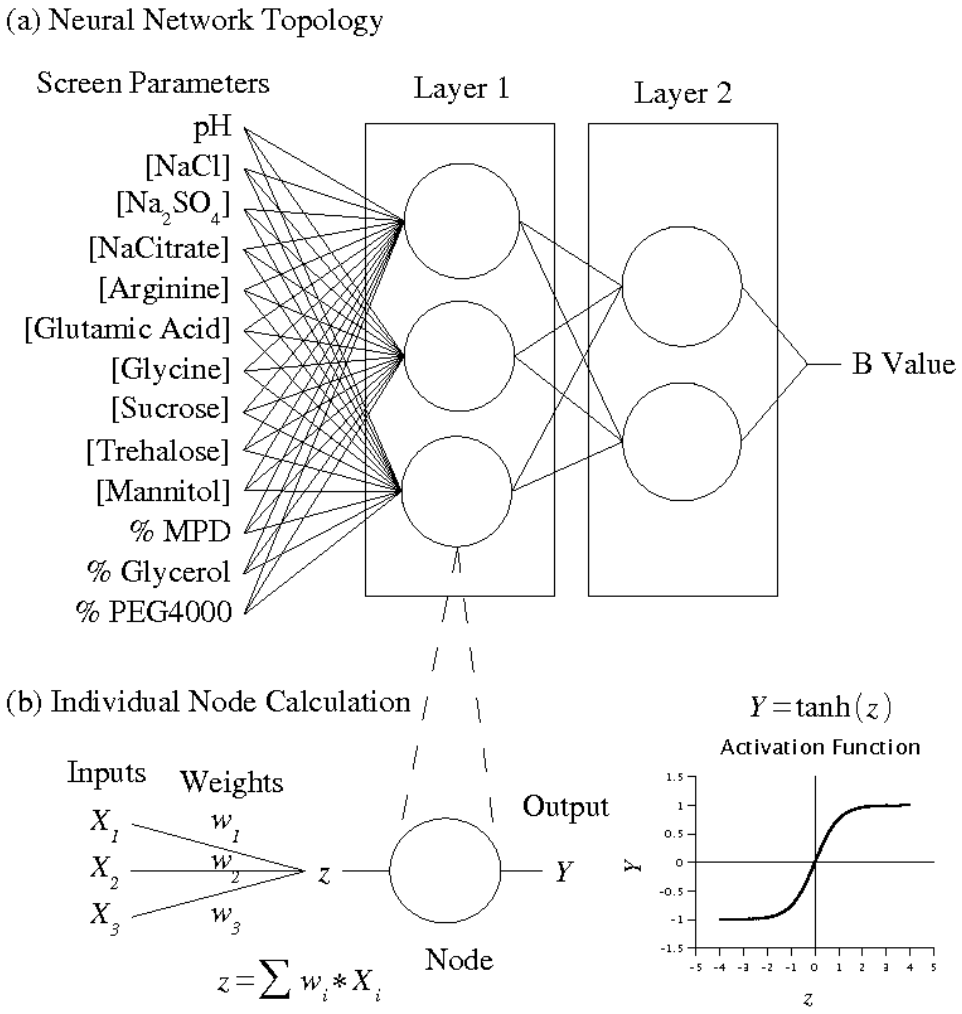

Figure 2. The artificial neural network topology (a) uses parameters of a single formulation as input to each node in Layer 1. Each node's output (b) is calculated by an activation function (tanh) whose input is a weighted sum of the node input. The output of nodes in layer one are forwarded as the input to Layer 2. The output of nodes in Layer 2 are weighted and summed to produce a B value prediction based on the input formulations.

This architecture (input vector, layers, and output) is generally referred to as a feed-forward multilayer perceptron and is capable of modeling a continuous function to arbitrary accuracy given a sufficient number of nodes[27]. Arbitrary accuracy is apparent if one considers a network topology containing one node for each formulation condition (N=81). After training the weight parameters, the response of each node could represent

the measured B value for each specific formulation condition. Such an exact fit to the screen would result in over-fitting to the error inherent to the screen and would therefore not provide a good generalized response to formulation conditions outside the screen on which it is trained.

To address the problem of over-fitting we split the set of screen conditions into a training set (90%) and a validation set (10%). During training the weights are iteratively adjusted using the gradient decent algorithm of back-propagation. This algorithm assigns an error contribution and updates each weight based on the root mean square error (RMSE) between the neural network output and the measured second virial coefficient for each formulation condition in the training set. RMSE is also calculated between the neural network output and measured B values in the validation set for each iteration. This validation RMSE is not used to improve weight values, but instead acts as the basis for deciding when to terminate the training procedure. The network weights are fixed at the minimum validation RMSE over a fixed number of iterations (1000). Validation set RMSE is also used as a measure of how well a network topology is able to generalize to untested formulation conditions. All network topologies from 1x1 to 6x6 nodes were evaluated by a validation set RMSE to determine the 3x2 network topology used for this study. Further details about neural network algorithms and methods can be found in Bishop's review[27] of the subject as well as in the JOONE software documentation[32].

Stepwise Generalized Linear Model (GLM)

The stepwise generalized linear model was performed using the JMP[33] statistical software package. The neural network inputs shown in Figure 2 were also the parameters used for the GLM. However, with GLM interaction and higher order terms must be explicitly identified for consideration. In addition to the neural network inputs, all pairwise interactions and square terms of the formulation screen were included for consideration. The stepwise algorithm was configured to include terms with a significance of alpha < 0.20 with higher order and interaction terms restricted to only those whose lower order terms were also significant.

Prediction Verification

The second virial coefficient for all combinations of buffer, salt and a maximum of two excipients (12,636 conditions) were predicted by the trained ANN. Five conditions from the most positive B values and five of the most negative B values as well as ten equally spaced throughout the range of predicted B values were selected for experimental confirmation. These 20 verification formulations (not included in the training process) were prepared and B values experimentally measured using the identical method as the original 81 screen conditions.

The question of whether 81 screen conditions are necessary or whether a smaller subset would suffice was addressed by evaluating the ability of the neural network to predict the verification B values while training on a reduced set of the initial screen. First a condition was randomly removed from the original training set of the neural network.

The training process described above was repeated on the reduced training set with the same validation set size remaining constant (deemed a valid indication of the overall population). Then the neural network, trained on a reduced set of the original 81 condition screen, was used to calculate predictions for the verification B values. Progressively reducing the sample size, followed by training and prediction, allows error as a function of sample size to be evaluated.

To determine how sample size affects neural network B value predictions, the validation set was kept constant while iteratively removing a random condition from the training set. As there is no consensus in the literature as to how this type of analysis should be performed a constant validation set was chosen as a good measure of the ability for the network to generalize. In keeping the same validation set through repeated reductions in the training set size we are able to see how available training data influences accuracy.

Results and Discussion

Confirmation by Static Light Scattering

A strong correlation (r=0.97) between static light scattering and self-interaction chromatography was observed (Table II) for ten test conditions as has been previously reported by other laboratories[20,28]. The primary differences between the two measurements were found with very positive values. Values in this range were expected to exhibit greater error since large positive B values have been shown to correspond to very high levels of solubility[1, 6]. Thus a small difference in B value represents a larger

difference in solubility. Therefore, from a practical perspective all high positive B values represent regions of high protein solubility even though individual B value errors are larger in this region.

Table 2. Comparison between B values measured by static light scattering (SLS) and self-interaction chromatography (SIC).

Condition ID	SLS B[a]	SIC B[a]
9	9	9.3
24	-1.4	-1.3
27	14	11
35	9	5.3
36	1.4	2.1
39	-1.4	-1.3
46	-0.5	0.0
60	7	7.7
72	3.7	3.8
79	-5	-3.6

[a] measured in * 10^{-4} mol ml / g^2

Screen Results

The full screen of 81 formulation conditions demonstrates some characteristics expected of lysozyme. For example, the mean B of the screen is positive 1.1 which is reflective of the general soluble nature of lysozyme. Additionally, a majority of the formulation conditions (55%) reside in the crystallization slot identified by George and Wilson[3] (George) which is approximately [-8,-0.2] x 10^{-4} mol ml/g^2. This is indicative of the ease with which lysozyme crystals are formed. It is also of interest to note that the average standard deviation between measurements was 1.7 x 10^{-4} mol ml/g^2. This

suggests that B measurements produced using self-interaction chromatography are reproducible throughout a large range of different solution conditions.

Interesting trends are also observed when viewing the influence of a single parameter throughout the screen. Figure 3 shows a graph of B value vs three individual parameter concentrations. The variation between plotted B values at a fixed concentration is due to the fact that other additives change with each condition. Error bars around each point indicate the error from measurement to measurement for each specific formulation. The increasing lysozyme self interaction (decreasing B) with increased concentration of sodium chloride (Figure 3a) is expected and has been demonstrated in other studies by both SIC and SLS[28]. At the mid and high concentrations of NaCl, four of the five conditions with positive B values contain MPD. This combined with the fact that MPD shows a trend (Figure 3b) of decreasing lysozyme self interaction (increasing B) with increasing concentration identifies MPD as a potential solubilizing agent for lysozyme. Quadratic relationships between additive concentration and B value, such as that apparent in glycine (Figure 3c) could also indicate an additive which might help stabilize protein self-interaction at a specific level. These single factor cross sections are useful for identifying individual additives which have a strong influence on B value. However, the prediction capability of single variable linear and quadratic regression models is obviously not sufficient to capture the variability in protein-protein interactions caused by formulations with multiple additives.

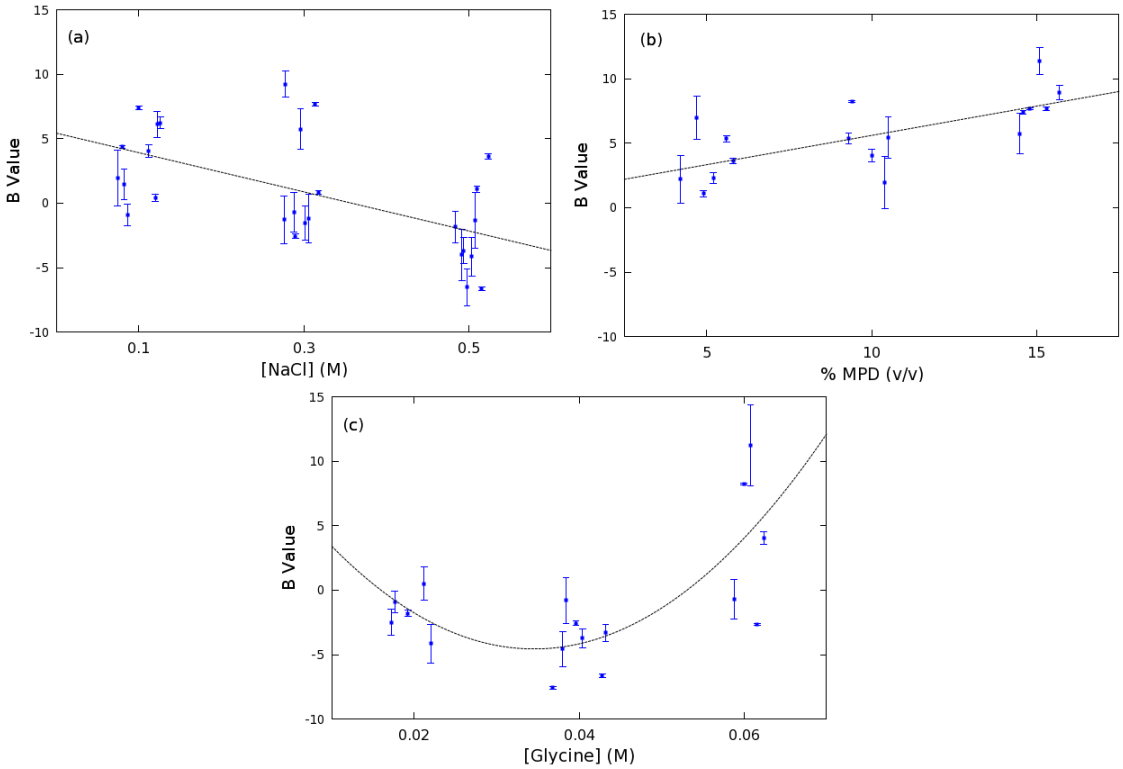

Figure 3. Response of B value for lysozyme by (a) NaCl (F-test;df=1;p=0.0006) , (b) MPD (F-test;df=1;p=0.001) and (c) Glycine (F-test;df=2;p=0.006) throughout all screen conditions containing the additive of interest. Error bars represent standard error between SIC measurements between whereas variability between points at a fixed additive concentration is attributed to changes in formulation parameters outside the additive of interest. Scatter along the abscissa is added to prevent overlapping of error bars.

Modeling and Prediction Results

The neural network trained on all conditions, except for 9 (10%) reserved for validation, produces a model which predicts the original screen with a RMSE of 1.7 x 10^{-4} mol ml/g^2. This is equal to the observed average standard deviation between measured B values and reinforces the notion that early termination of training based on the validation set error prevents over-fitting of the screen. Upon completion of training, the neural network is used to predict B values for all possible variable combinations with one or two additives (12,636 formulation conditions). From this entire number of predictions,

twenty predictions were chosen for verification. These twenty conditions were chosen to represent the entire solubility range, with some from the most positive and negative predicted B values. These formulation conditions and their predicted second virial coefficients are shown in Table III. The experimental formulations in Table III were prepared and their effect on lysozyme's second virial coefficient measured via SIC. The plot of measured B values versus ANN predicted values in Figure 4 demonstrates that the neural network is able to predict second virial coefficients with an accuracy of 2.6×10^{-4} mol ml/g^2.

Table 3. ANN Predictions of 20 formulations selected for verification.

Buffer	Salt	Excipient 1	Excipient 2	Predicted B[a]
0.1M HEPES	0.5M NaCl	0.04M Glycine	0.04M Arginine	-6.0
0.1M Succinate	0.5M Na_2SO_4	0.06M Glycine	0.2M Mannitol	-5.8
0.1M HEPES	0.5M NaCl	0.04M Glutamic Acid	0.3M Mannitol	-5.5
0.1M Acetate	0.5M Na_2SO_4	0.06M Glycine	0.1M Sucrose	-5.1
0.1M MES	0.5M NaCl	0.3M Mannitol	0.2M Trehalose	-5.1
0.1M Succinate	0.5M NaCl	0.06M Glycine	0.3M Sucrose	-3.6
0.1M HEPES	0.5M Na_2SO_4	0.06M Arginine	0.1M Mannitol	-2.6
0.1M HEPES	0.1M Na_2SO_4	0.02M Arginine	0.1M Mannitol	-1.8
0.1M Succinate	0.5M NaCl	0.3M Trehalose	5% MPD	-1.1
0.1M Succinate	0.5M NaCitrate	0.02M Glycine	0.3M Sucrose	-0.4
0.1M Acetate	0.3M Na_2SO_4	0.06M Arginine	0.2M Trehalose	0.4
0.1M Acetate	0.1M NaCl	0.06M Arginine	-	1.5
0.1M Succinate	0.1M Na_2SO_4	0.2M Mannitol	10% PEG4000	2.8
0.1M MES	0.3M NaCl	15% MES	-	4.4
0.1M HEPES	0.3M NaCl	10% MPD	9% Glycerol	6.8
0.1M MES	0.3M NaCitrate	9% Glycerol	10% PEG4000	9.1
0.1M MES	0.3 NaCl	0.04M Glycine	15% PEG4000	9.3

0.1M MES	0.1M Na_2SO_4	0.06M Glycine	15% MPD	9.8
0.1M HEPES	0.3M NaCitrate	0.1M Trehalose	15% MPD	13
0.1M HEPES	0.3M NaCitrate	10% MPD	10% PEG4000	14

[a] measured in * 10^{-4} mol ml / g^2

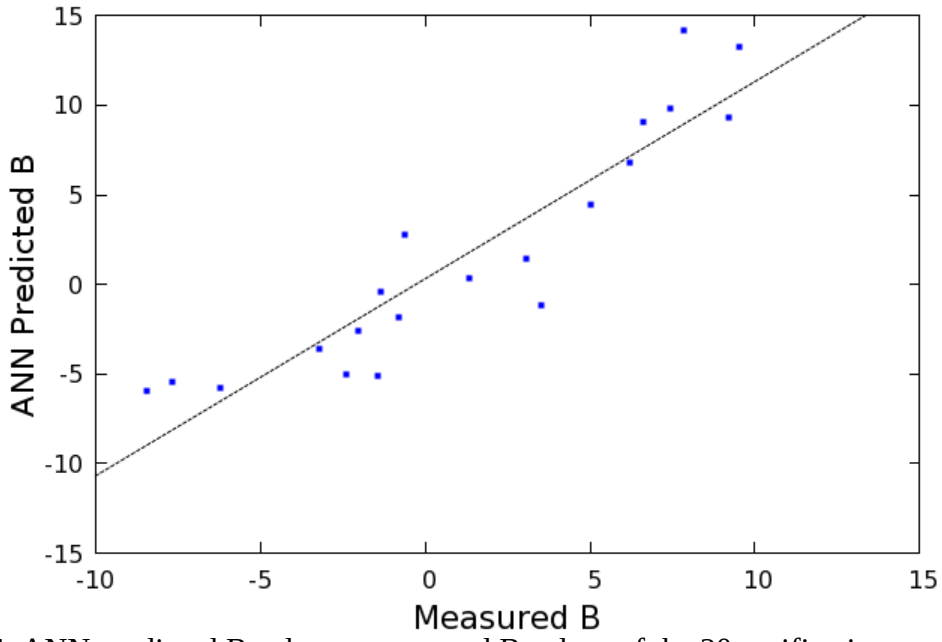

Figure 4. ANN predicted B value vs measured B values of the 20 verification formulations (F-test;df=1;p<0.0001;RMSE=2.6 x 10^{-4} mol ml/g^2).

Sample size plays a role in how accurately the ANN model is able to predict untested formulation conditions. Figure 5 shows that the prediction error of ANN increases as the size of the available training set decreases. Although prediction error increases as training size decreases an error of 3 B units is still attainable with a sample size of 45 screen conditions. Also, while prediction error decreases as training size increases the error reduction appears to diminish with a larger sample size. However the curve does not completely flatten suggesting a screen size of over 100 conditions could permit ANN prediction with an error closer to 2 B units.

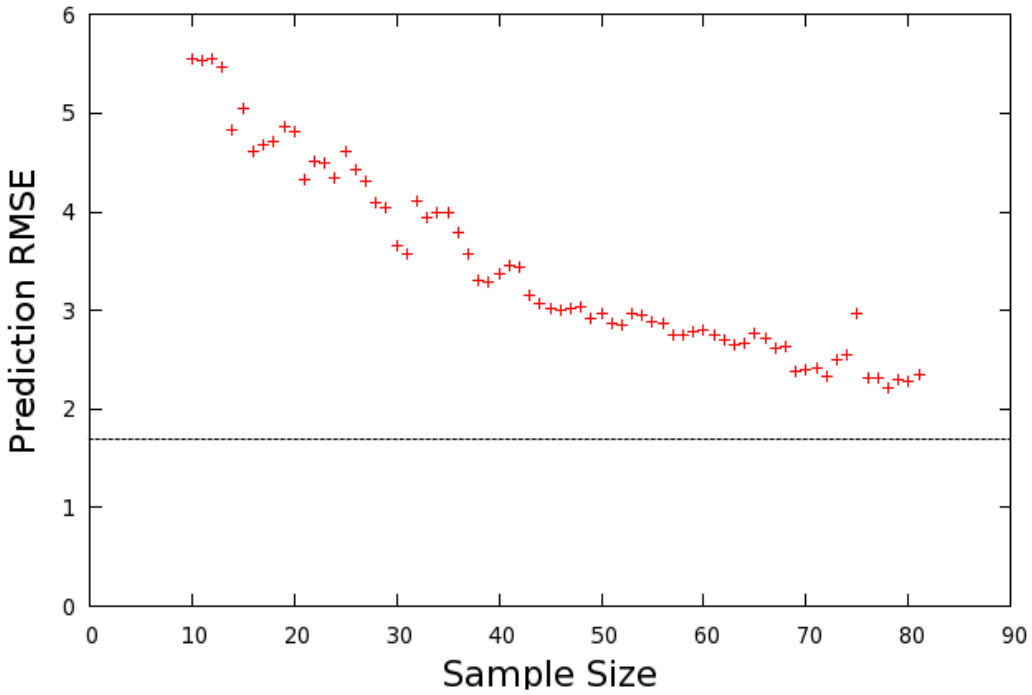

Figure 5. ANN RMSE vs sample size. Incremental reduction in sample size shows an increase in error for artificial neural network predictions of the 20 verification formulations. Dashed line indicates the error between B value measurements by SIC between columns (1.7 mol ml / g^2).

The standard generalized linear model provides a comparison of the ANN with a standard linear regression technique used for data analysis/predictions. The terms of the GLM were determined by considering all single terms, interaction terms and square terms and incrementally adding the most significant remaining parameter until there are no more parameters with a significance of alpha < 0.20. The GLM parameters and their significance level generated by this method are listed in Table IV. This table demonstrates one benefit of the GLM over ANN. Incremental analysis of each parameter produces a list of significant factors. This helps identify specific formulation parameters which could increase solubility. However, when predicting the second virial coefficient of protein in previously unformulated conditions the GLM does not perform as well as the ANN. The plot in Figure 6 shows the same 20 measured B values for ANN

validation versus the GLM predictions. Although both predictions are statistically significant (F-test;df=1;p<0.0001), the GLM is accurate with a RMSE of 3.3×10^{-4} mol ml/g^2 which implies the ANN is approximately 25% more accurate than the GLM. However both techniques are be useful for formulation prediction based on a small subset of conditions.

Table 4. Additives with statistically significant influence as determined by stepwise GLM.

Factor	p-value	Magnitude
Glycine	< 0.0001	-1.5
MPD	< 0.0001	2.5
NaCl	< 0.0001	1.4
PEG4000	< 0.0001	2.1
Arginine	0.0001	-1.3
Citrate	0.0004	0.5
Mannitol	0.0127	-0.6
Glycerol	0.0562	-1.4
Glycine * PEG4000	< 0.0001	1.7
Arginine * Glycine	0.0002	-2.2
NaCl * MPD	0.0099	-0.5
NaCl * Mannitol	0.0154	-0.4
Citrate * Glycine	0.0272	0.3
Citrate * Mannitol	0.1573	-0.2
Glycine * Glycerol	0.1685	-0.5
$NaCl^2$	< 0.0001	-0.7
$Glycerol^2$	0.0108	1.0

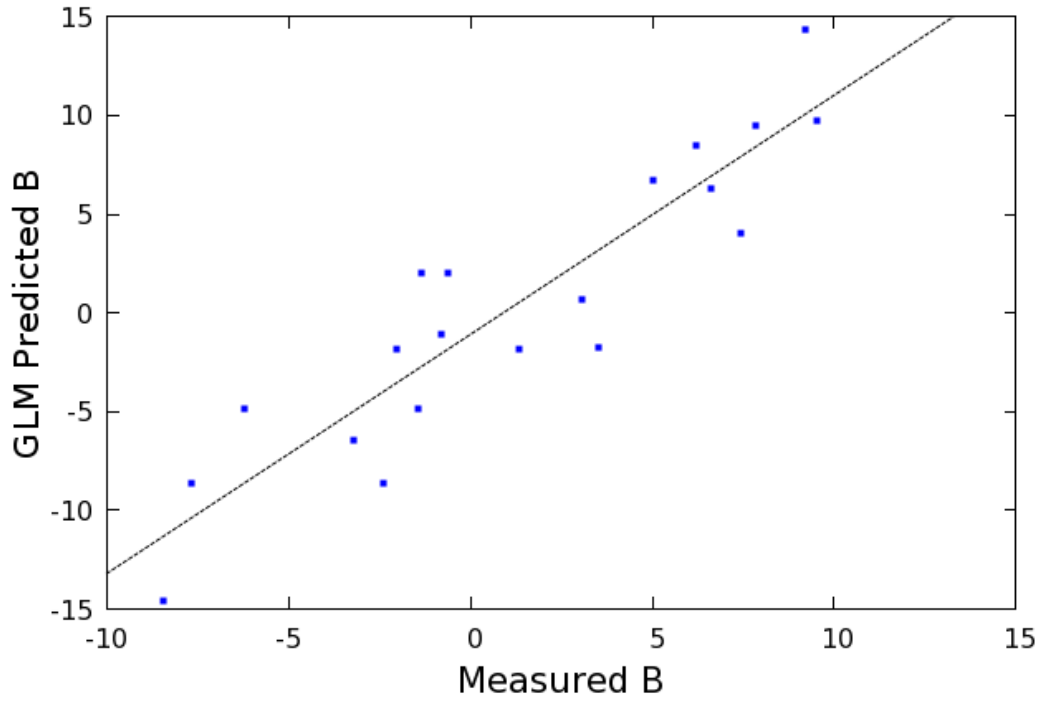

Figure 6. GLM predicted B values vs measured B values of the 20 verification formulations (F-test;df=1;p<0.0001;RMSE=3.3 x 10^{-4} mol ml/g^2).

Limitation

A limitation of this screen and formulation prediction technique is in the ability to predict formulation conditions with parameter concentrations well outside the screened range. The inability for statistical models to extrapolate results outside their original input range is well known. This implies that the range of pH and salt/additives concentrations must be chosen based on an estimation of the effective range for each parameter. For example the pH range of interest might be a region in relation to the expected pI of the protein. It is important to note that once parameter ranges are determined the screen and resulting statistical models will not be able to predict the B

value of formulations with parameters significantly outside these ranges. However, this does not diminish the fact that the statistical models can accurately predict the B value of a large number of novel formulation conditions based on parameter combinations not measured in the original screen.

Conclusions

As hypothesized in previous publications[9,21,25], high throughput screening of second virial coefficients shows promise for evaluating the interactions of proteins in solution. We have demonstrated that an incomplete factorial screen combined with a neural network model can be used to accurately predict second virial coefficients for untested formulations. A B value screen of only 81 formulation conditions was used to predict the B values for 12,636 possible formulations with an accuracy of $2.6 * 10^{-4}$ mol ml / g^2. These preliminary studies suggest that a high-throughput chromatographic SIC system with increased automation may enhance and accelerate determinations of the optimum conditions that improve the physical solubility/stability of drug formulations. It also suggests that the same technology may be useful to predict formulation adjustments required for optimized protein expression and/or crystallization.

The current time required to run self-interaction chromatography in triplicate is approximately 30 minutes. While 30 minutes per experiment by SIC is much faster than previous SLS methods[20], the use of B values for these applications would benefit significantly by increased throughput via parallelization, robotic automation and integration of analysis techniques into a single platform.

Acknowledgements

This work was supported by a grant from the Alabama EPSCoR Graduate Student Research Scholars program. Thank you to Dr. Lisa Nagy for training on the Shimadzu HPLC hardware.

References

1. Haas, C., Drenth, J. & Wilson, W. Relation between the Solubility of Proteins in Aqueous Solutions and the Second Virial Coefficient of the Solution. *J. Phys. Chem. B* **103**, 2808-2811(1999).
2. Ahamed, T. et al. Design of self-interaction chromatography as an analytical tool for predicting protein phase behavior. *Journal of Chromatography A* **1089**, 111-124(2005).
3. George, A. & Wilson, W.W. Predicting Protein Crystallization from a Dilute Solution Property. *Acta Crystallographica Section D* **50**, 361-365(1994).
4. Neal, B.L., Asthagiri, D. & Lenhoff, A.M. Molecular Origins of Osmotic Second Virial Coefficients of Proteins. *Biophys. J.* **75**, 2469-2477(1998).
5. McMillan, J. & Mayer, J.E. The Statistical Thermodynamics of Multicomponent Systems. *J. Chem. Phys.* **13**, 276-305(1945).
6. Demoruelle, K. et al. Correlation between the osmotic second virial coefficient and solubility for equine serum albumin and ovalbumin. *Acta Crystallogr. D Biol. Crystallogr* **58**, 1544-8(2002).
7. Chi, E.Y. et al. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. *Pharmaceutical Research* **20**, 1325-1336(2003).
8. Guo B. et al. Correlation of second virial coefficients and solubilities useful in protein crystal growth. *Journal of Crystal Growth* **196**, 424-433(1999).
9. Tessier, P.M. et al. Self-interaction chromatography: a novel screening method for rational protein crystallization. *Acta Crystallogr. D Biol. Crystallogr* **58**, 1531-5(2002).

10. Valente, J.J. et al. Colloidal Behavior of Proteins: Effects of the Second Virial Coefficient on Solubility, Crystallization and Aggregation of Proteins in Aqueous Solution. *Current Pharmaceutical Biotechnology* **6**, 427-436(2005).

11. Christendat, D. et al. Structural proteomics of an archaeon. *Nat Struct Mol Biol* **7**, 903-909(2000).

12. Luan, C. et al. High-Throughput Expression of C. elegans Proteins. *Genome Res.* **14**, 2102-2110(2004).

13. Singh, S.M. & Panda, A.K. Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering* **99**, 303-310(2005).

14. Koo, E.H., Lansbury, P.T. & Kelly, J.W. Amyloid diseases: Abnormal protein aggregation in neurodegeneration. *Proceedings of the National Academy of Sciences* **96**, 9989-9990(1999).

15. Pande, A. et al. Crystal cataracts: Human genetic cataract caused by protein crystallization. *Proceedings of the National Academy of Sciences* **98**, 6116-6120(2001).

16. Ho, J.G. et al. The likelihood of aggregation during protein renaturation can be assessed using the second virial coefficient. *Protein Sci* **12**, 708-716(2003).

17. Loftsson, T. & Brewster, M.E. Physicochemical properties of water and its effect on drug delivery. A commentary. *International Journal of Pharmaceutics* **354**, 248-54(2008).

18. Kerns, E. & Di, L. Physicochemical profiling: overview of the screens. *Drug Discovery Today: Technologies* **1**, 343-348(2004).

19. Patro, S.Y. & Przybycien, T.M. Self-interaction chromatography: A tool for the study of protein-protein interactions in bioprocessing environments. *Biotechnology and Bioengineering* **52**, 193-203(1996).

20. Tessier, P.M., Lenhoff, A.M. & Sandler, S.I. Rapid measurement of protein osmotic second virial coefficients by self-interaction chromatography. *Biophys. J* **82**, 1620-31(2002).

21. Valente, J.J. et al. Screening for physical stability of a Pseudomonas amylase using self-interaction chromatography. *Anal. Biochem* **357**, 35-42(2006).

22. Carter, C.W. & Carter, C.W. Protein crystallization using incomplete factorial experiments. *J. Biol. Chem* **254**, 12219-23(1979).

23. Taguchi, G. & Konishi, S. *Orthogonal arrays and linear graphs*. (American Supplier Institute Dearborn, Mich: 1987).

24. Kingston, R.L., Baker, H.M. & Baker, E.N. Search Designs for Protein Crystallization Based on Orthogonal Arrays. *Acta Crystallographica Section D* **50**, 429-440(1994).

25. DeLucas, L.J. et al. Protein crystallization: virtual screening and optimization. *Progress in Biophysics and Molecular Biology* **88**, 285-309(2005).

26. Souza, L.E.S.D. & Canuto, S. Efficient estimation of second virial coefficients of fused hard-sphere molecules by an artificial neural network. *Phys. Chem. Chem. Phys.* **3**, 4762-4768(2001).

27. Bishop, C.M. Neural networks and their applications. *Rev. Sci. Instrum.* **65**, 1803-1832(1994).

28. Valente, J.J. et al. Second Virial Coefficient Studies of Cosolvent-Induced Protein Self-Interaction. *Biophys. J.* **89**, 4211-4218(2005).

29. DePhillips, P. & Lenhoff, A.M. Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography. *Journal of Chromatography A* **883**, 39-54(2000).

30. Fanguy, J. et al. Scale-Down Approaches for Measuring Protein-Protein Interactions. *Protein Crystallization Strategies for Structural Genomics* 127-152(2007).

31. Kratochvil, P. *Classical light scattering from polymer solutions*. (Amsterdam; New York: Elsevier; New York, NY, USA: Distributers for the US and Canada, Elsevier Science Pub. Co.: 1987).

32. Marrone, P. Joone - Java Object Oriented Neural Engine. at <http://www.jooneworld.com/>

33. *JMP*. SAS Institute Inc., Cary, NC, 1989-2007.

Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulation Equipment

David H. Johnson, Martha W. Bidez, Lawrence J. DeLucas

Annals of Biomedical Engineering **40,** 898–906 (2012).

Abstract

Hazard analysis and risk assessment techniques are utilized within many private sector industries and government agencies, including the medical device and pharmaceutical industry, within a structured process to control human injuries and environmental and property damage. In the U.S. the Federal Drug Administration (FDA) requires a hazard analysis be performed on all medical devices. While there are biomedical engineering applications reported which deal with human hazards in clinical, patient care environment, no previous studies extend these traditional techniques to a university-based, research environment. This study applies a tiered approach to hazard analysis and risk assessment to a biomedical, university-based, research environment in the design of a high throughput platform that screens chemical excipients (additives) for their ability to increase protein solubility. Each design stage (conceptual, preliminary, system and detailed) requires a unique hazard analysis technique based on available information. The analysis techniques applied here are evaluated for their use in a biomedical research environment where experiment accuracy is a primary concern.

Key Terms: hazard analysis, risk assessment, risk reduction, failure mode effects analysis, ANSI, GEIA, self-interaction chromatography

Introduction

Hazard analysis and risk assessment techniques are utilized within many private sector industries and government agencies, including the medical device and pharmaceutical industries[1,5,8]. A hazard is most generally defined as a potential source of harm[1] and more specifically, a source of physical harm to people, property and/or the environment [7]. A hazard may also refer to an interruption in function or operations. In hazard analysis, potential sources of physical and/or functional harm are identified, prioritized and mitigated in a structured manner. The practical goal of hazard analysis is to identify ways to reduce risk to an acceptable level where the next incremental reduction in risk potential is not justified by the cost of further risk reduction [7]. To determine this threshold for deciding against further action, the potential risk associated with each hazard must be identified and analyzed – a process known as "risk assessment." The risk associated with a given hazard is a function of both the severity of potential outcome from the hazard, as well as, the expected probability that an incident (harm) will occur. Risk assessment focused on the probability of failures is well established in probabilistic risk assessment and fault tree analysis. Thus, risk assessment based on objective analysis of hazard severity and probability of occurrence can provide an important guide to determine the amount of resources that should be effectively invested in reducing a given risk, as well as, identifying unacceptable risk.

The risk assessment process involves four discreet steps over the entire lifecycle of a system after program initiation (Figure 1). These include (1) hazard identification, (2) risk assessment, (3) risk reduction and (4) risk acceptance. The process also requires continuous hazard tracking.

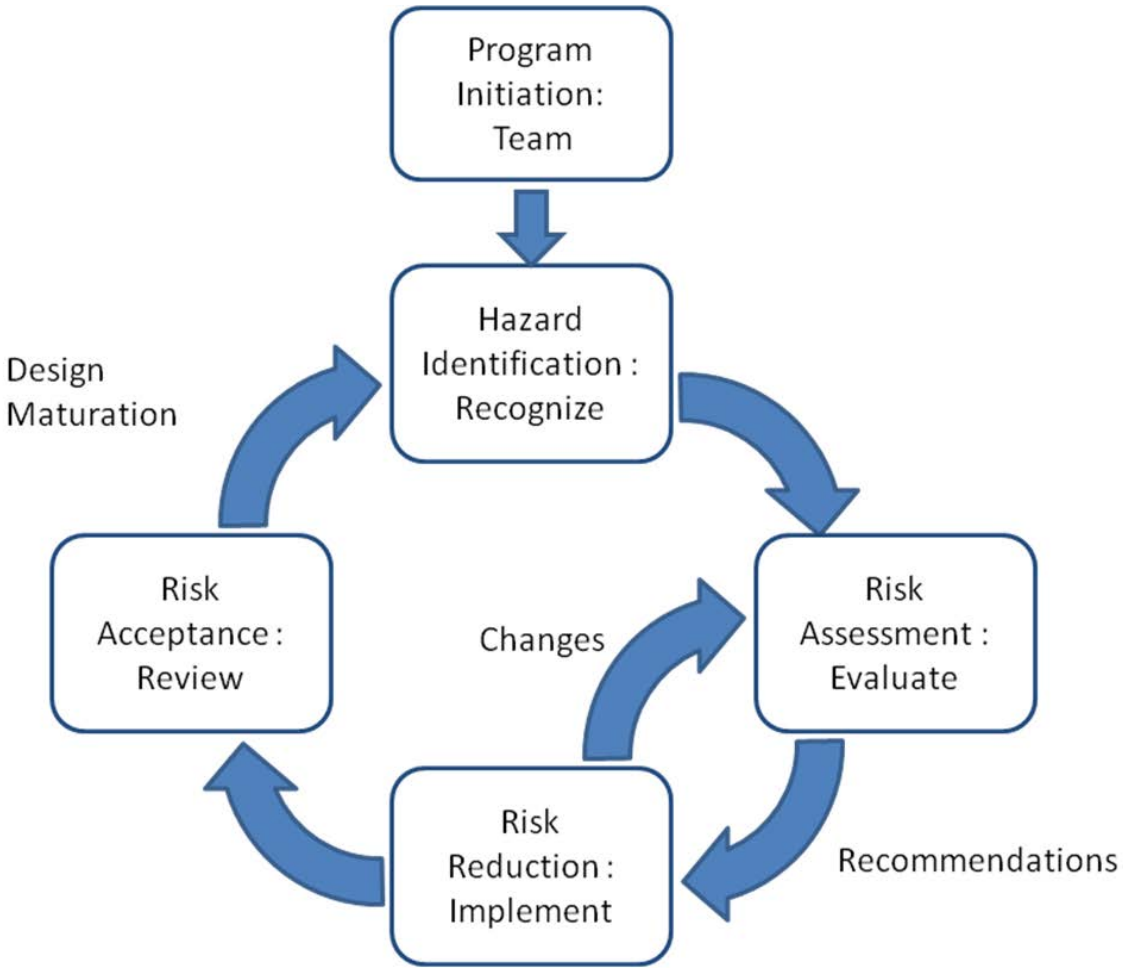

FIGURE 1: The Risk Assessment Process (Adapted from ANSI/GEIA-STD-0010[1])

Identifying hazards is the first step in the risk assessment process. The field of system safety relies upon seven basic hazard analysis types, which provide the framework for over one hundred different hazard analysis techniques. [Ericson[4], p.31] A hazard analysis type establishes a specific analysis task at a specific time in the program or system life cycle, whereas, a hazard technique establishes a specific and unique analysis methodology. The seven basic hazard analysis types and common analysis

techniques associated with those types are provided in Table 1. A comprehensive review of the benefits and limitations of each technique can be found in Ericson[4].

Table 1: Common Hazard Analysis Types and Techniques (Ericson, p. 48)

Primary Analysis Technique	Hazard Analysis Type
Preliminary Hazard List (PHL)	Conceptual Design Hazard Analysis Type (CD-HAT) *Purpose is to compile a list of hazards very early in the system development life cycle*
Preliminary Hazard Analysis (PHA)	Preliminary Design Hazard Analysis Type (PD-HAT) *Purpose is to identify system-level hazards and to obtain initial risk assessment of a system design*
Failure Mode & Effects Analysis (FMEA); Hazard and Operability Analysis (HAZOP)	Detailed Design Hazard Analysis Type (DD-HAT) *Purpose is to identify single-point failures in an analysis of the detailed design.*
Functional Hazard Analysis (FuHAT); Fault Tree Analysis (FTA); Event Tree Analysis (ETA)	System Design Hazard Analysis Type (SD-HAT) *Purpose is to evaluate the integrated system design and identify subsystem interface hazards*
Operations & Support Hazard Analysis (O&SHA)	Operations Design Hazard (OD-HAT) *Purpose is to evaluate operations and support functions involved with the system, including consideration of human limitations*
Health Hazard Assessment (HHA)	Health Design Hazard Analysis (HD-HAT) *Purpose is to identify human health hazards*
Safety Requirements/Criteria Analysis (SRCA)	Requirements Design Hazard Analysis (RD-HAT) *Purpose is to verify and validate the design safety requirements and ensure that no safety gaps exist in written requirements*

The time period for performing specific hazard analysis techniques in an engineering development life-cycle is illustrated in Figure 2. The timeframe for three significant design review meetings is also shown in Figure 2.

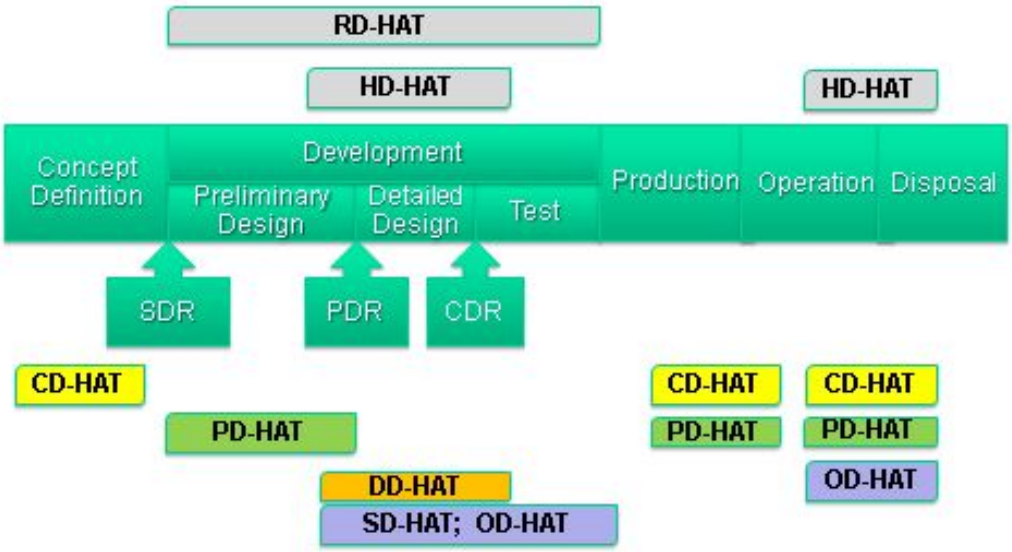

Figure 2: Overall timing of hazard analysis types in the life cycle of an engineering design.

The Functional Hazard Analysis (FuHA) technique, a system design hazard analysis type, is particularly well suited for use in the laboratory research environment. The primary purpose of FuHA is to identify hazards through an analysis of a system's functions. Outcomes from a FuHA include the consequences of a system function, which fails to operate, operates incorrectly and/or operates at the wrong time. A safety critical function is defined as "any function whose failure or misbehavior could result in death/injury and/or system loss." [Ericson[4], p. 478] While FuHA is common in aerospace

and military applications, the technique is mentioned in neither ISO international standards, nor FDA regulations and guidance documents. To the authors' knowledge, the FuHA has not been previously applied in the field of biomedical engineering, in general, and research laboratories, in particular. Functional hazard analysis can provide a tool for structured identification and robust evaluation of those hazards (experimental variables and/or design issues), which can adversely affect the validity of research findings.

Purpose

The purpose of this study was to conduct a risk assessment on a novel, custom-designed system for the development of biomedical drug formulations and in doing so, demonstrate the applicability and utility of functional hazard analysis in biomedical engineering design.

Methods

In this study, we assessed the risks associated with a custom-designed, high throughput platform designed to screen chemical excipients (additives) for their ability to increase protein solubility ("System"). This system involves a university-based, biomedical research laboratory, which specializes in the development and optimization of protein drugs, which play a role in many disease therapies, including insulin and antibody-targeted chemotherapy. A brief review of the scientific principles in this

research is necessary in order to appreciate the need for and ultimate conclusions reached from a comprehensive hazard analysis and risk assessment process.

A key component of drug design is an understanding of the interactions between protein molecules. These interactions between proteins are critical to both the solubility of proteins and their abililty to crystallize. Slight attraction between protein molecules is required for protein molecules to gently come together to form a crystal lattice and when combined with xray diffraction analysis, provides the most accurate way to determine the structure of protein molecules. Knowledge about a protein structure is useful for the determination of that protein's function in the body and for the understanding of how protein defects can cause disease states in humans. While slightly attractive forces between protein molecules are necessary for crystal formation, repulsive forces between molecules are associated with increased solubility. Protein drug therapies that aggregate can reduce the efficacy of the drug and potentially cause an immune response in patients[11]. Therefore, the U.S. Food and Drug Administration (FDA) requires protein drugs to be soluble and physically stable at a high concentration. The identification of formulations in which protein molecules repel each other is a critical step to produce safe and usable protein drug therapies.

A novel design for a high-throughput self-interaction chromatography (HSC) machine was conceived and developed by the authors to rapidly determine the influence of chemical formulation excipients on protein-protein interactions. This information is used to identify high solubility or slightly insoluble formulations depending on whether or not the application is protein drug therapeutics or structure studies by x-ray crystallography, respectively. Excipient influence on protein self-interaction is

established in two stages: the excipient screen and excipient influence analysis. During the screening process each excipient formulation is sequentially delivered to a custom, four-channel, liquid chromatography system. The conception and development of the hardware and software to measure protein-protein interactions in a variety of formulation conditions was defined as the system for this hazard analysis and risk assessment.

A schematic diagram of the formulation flow system is illustrated in Figure 3. The main formulation pump drives the formulation to be analyzed through the system. When withdrawing a formulation, the formulation valve is switched to the reservoir tray located on the formulation platform (Figure 4). There is an option valve which can deliver the formulation through either an acetone (control) or protein injection valve depending on which fluid is to be injected. When a protein is injected, the protein injection valve switches the flow from the formulation pump through the column to the protein pump through the column. The same procedure occurs for injection of acetone using the acetone injection valve and acetone pump. After injection the injection pump switches back to allow the formulation pump to push formulation (with a bolus of protein or acetone) through the system. The injected volume flows over the column, and is detected by UV absorbance after the column. The time required to push the protein through the column compared to the non-interacting marker at the same flow rate is used to calculate the B value. Longer retention times of the protein compared to the non-interacting marker correspond to greater protein-protein interactions and more negative B values.

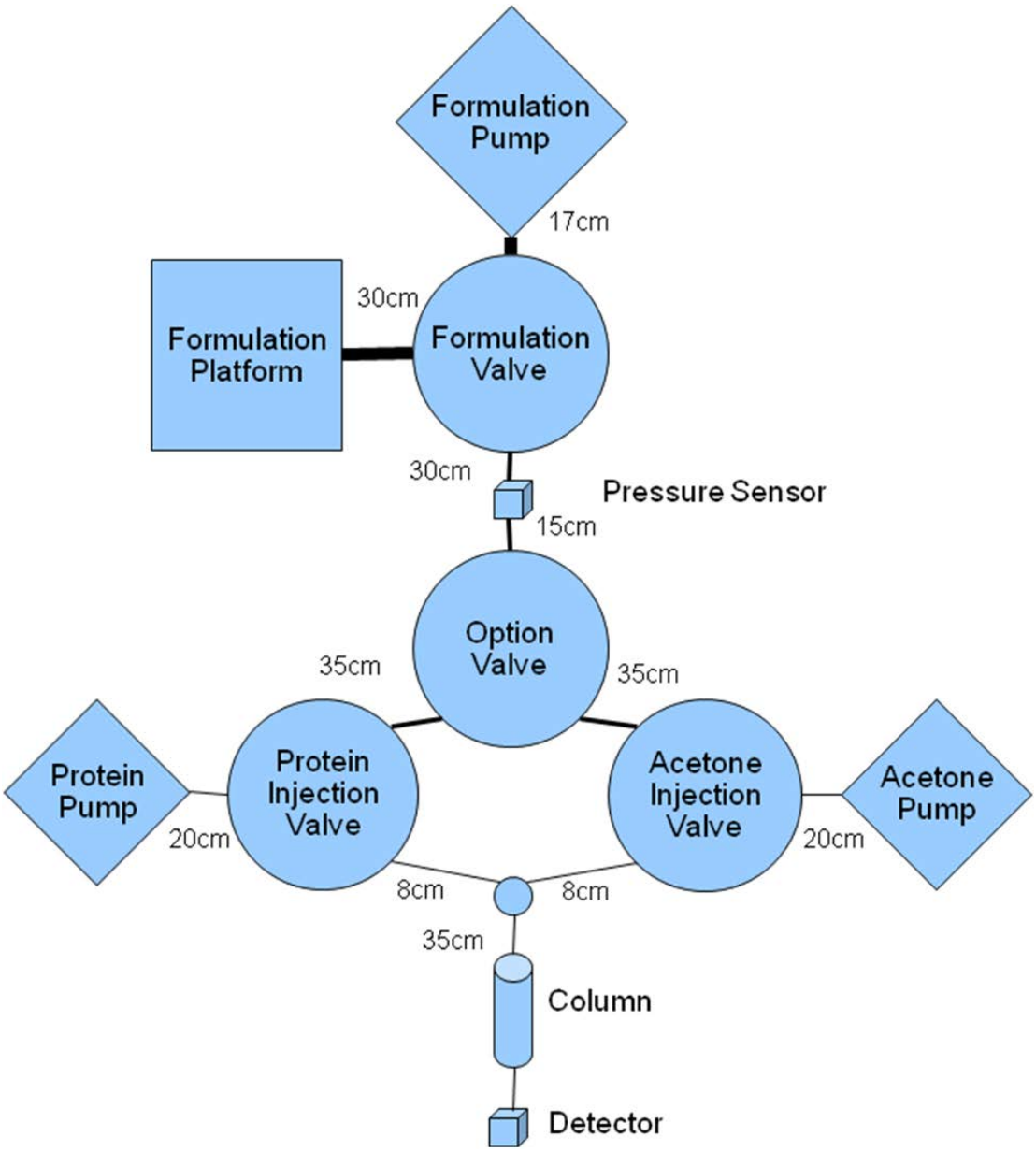

Figure 3: High-Throughput Self-Interaction Chromatography (HSC) flow diagram

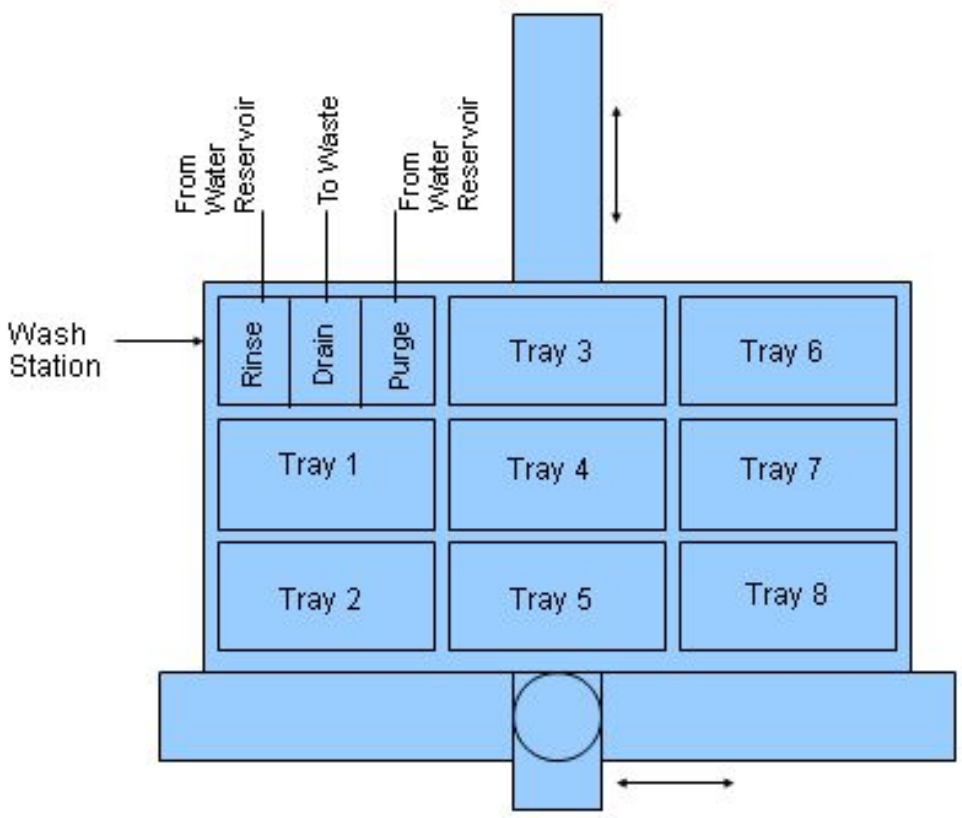

Figure 4: Formulation platform diagram

Four hazard analysis techniques were used in conducting the system risk assessment: (1) Preliminary Hazard List (PHL), (2) Preliminary Hazard Analysis (PHA), (3) Failure Mode Effect Analysis (FMEA), and (4) Functional Hazard Analysis. These four techniques corresponded to the Conceptual Design, Preliminary Design, Detailed Design and System Design hazard analysis types, respectively. They were chosen to correspond to the current state of the System; the Concept and Development stages of the system life cycle are completed (refer to Figure 2). The detailed methodology for the PHL, PHA and FMEA techniques are well accepted and published in the biomedical

literature and standards[1-10,12,13]. The methodology associated with the Functional Hazard Analysis technique is provided in Table 2.

Table 2: Functional Hazard Analysis Methodology (Ericson, p 274)

Step	Task	Description
1	Define operation	Define the scope and boundaries of the operation under study. In this study, the scope was both hardware and software associated with the system.
2	Acquire data	Acquire detailed operational data for the system, which includes schematics, walkthroughs, manuals, standards, etc. In this study, all of the above were utilized for the FuHA
3	List functions	Create a detailed list of all functions to be considered in the FuHA. This study evaluated both hardware and software system functions.
4	Conduct FuHA	Evaluate the effect of each functional failure mode in the system function list.
5	Evaluate system risk	Identify the level of mishap associated with each of the hazards. This study utilized the risk assessment matrix provided in Table 3.
6	Identify safety critical functions	Identify those functions, which are considered safety critical. In this study, those functions relate to people, property and experimental accuracy
7	Recommend corrective action	Recommend corrective action.to eliminate or mitigate the hazards related to people, property and experimental accuracy
8	Monitor corrective action	Review design requirements to ensure that appropriate corrective action is being taken. In this study, corrective action was evaluated and taken at the preliminary design stage.
9	Track hazards	Record hazards in appropriate hazard tracking system.
10	Document FuHA	Document the FuHA process on worksheets.

Two approaches were used to identify preliminary hazards: a walkthrough of how the system is typically used and a review of mandatory, government standards. Regulations, which were reviewed, included the FDA's ISO 17941, Environmental Protection Agency's (EPA) Good Laboratory Practice Standards [5] and the Occupational

Safety and Health Administration's (OSHA) standard for Toxic and Hazardous Substances [9].

Table 3: Risk Assessment Matrix

FREQUENCY	SEVERITY			
	Catastrophic (1)	**Critical (2)**	**Marginal (3)**	**Negligible (4)**
Frequent (A)	1	2	4	7
Probable (B)	3	5	8	11
Unlikely (C)	6	9	12	14
Improbable (D)	10	13	15	16

According to ANSI/GEIA-STD-0010-2009, severity and probability measurements can be evaluated on either a subjective (qualitative) or quantitative scale[1]. A 4x4 quantitative risk assessment matrix was used for all analyses (Table 3). Severity and frequency scales utilized in this study reflect the potential harm to people, equipment and experimental accuracy (Tables 4 and 5). The human injury severity scale was based on OSHA reportability requirements (an injury requiring more than first aid). Hazard severity relative to equipment were defined on the basis of the replacement cost of a specific piece of damaged equipment with the minimum and maximum values established commensurate with the operational budget of our specific research environment. Hazards to experiment accuracy were identified using functional hazard analysis (FuHA) and based on the ability of the system to collect accurate data during a screen of 80 formulation conditions.

Table 4. Hazard severity to Humans, Equipment and Experiments

Category	**Human**	**Equipment**	**Experiment***
Catastrophic (1)	Serious disability, death	> $50,000	Invalidate > 32 conditions (40%)

Critical (2)	Multiple days lost, minor disability	< $50,000	Reduced resolution > 32 conditions (40%)
Marginal (3)	Up to a day lost, OSHA Reportable	< $5,000	Reduced resolution > 8 conditions (10%)
Negligible (4)	First aid only, not OSHA Reportable	< $500	Reduced resolution < 8 conditions (10%)

* Percent based on 80 conditions per screen.

Table 5. Hazard frequency in terms of system use.

Category	**Description**	**Experiment**
Frequent (A)	Frequently observed.	< 10 screens per incident
Probable (B)	Rarely observed, but expected.	> 10 screens per incident
Unlikely (C)	Not yet observed or observed very rarely*	> 100 screens per incident
Improbable (D)	Not yet observed. Unexpected.	Unexpected

*Reasonable to expect in system lifetime (15 years).

Risk reduction efforts were prioritized for each hazard based on the assigned risk category. Risk reduction techniques evaluated potential changes to both procedures and system design. Changes to the standard operating procedures were evaluated to produce a more robust evaluation of proteins prior to screening. Changes to the system design, which were evaluated included increased monitoring and fail-safe controls. In our research environment, it was determined that the most critical monetary risk was less significant than moderate risk of human injury. Thus, all risks to people above a low level were ranked as a higher priority than operational risks.

Results

The risk assessment results are presented according to the position of the hazard analysis type on the engineering design life cycle (Figure 2).

Conceptual Design Hazard Analysis – Preliminary Hazard List (PHL)

The preliminary hazard list is an initial list of hazards in the system, determined in the early conceptual design phase (Appendix 1). For the HSC system the PHL identified ten hazards to humans, one to the environment, three to equipment and thirteen to operations for a total of 27 hazards. A significant design decision was made at this point to split the flow path for protein and acetone to avoid contamination of the signal between protein and control experiments. Hazards identified in the PHL are tracked and used in subsequent design stages as the minimum hazards to be addressed with the expectation that additional hazards will be identified as the design progresses.

Preliminary Design Hazard Analysis – Preliminary Hazard Analysis (PHA)

The preliminary hazard analysis identified potential hazards to people, equipment, operations and the environment and their causes and effects (Appendix 2). Potential hazards identified for people include slip hazards, uncontrolled system pressure release and screen formulation chemical interactions. Characterization of the hazards depended upon both the severity of the hazard and probability of occurrence. In the discussion we describe the hazard characterization, risk assessment and actions identified to improve the safety and productivity of our laboratory environment. Applicability to other research and clinical evironments is also discussed later.

System Design Hazard Analysis – Functional Hazard Analysis (FuHA)

First, a list of thirty-seven system functions was enumerated using test case scenarios (Appendix 3). Each function was evaluated for failure, malfunction or corruption. The functional evaluation included both hardware and software functions during data collection of a full screen of formulation conditions.

Detailed Design Hazard Analysis – Failure Mode and Effect Analysis (FMEA)

FMEA, a detailed design tool, was used to evaluate specific modes of failure of subsystems and components which could affect system reliability. FMEA (Appendices 4-5) evaluated individual components to identify seventy-four failure modes, including protein precipitation in the system (operational) and syringe plunger impact in the syringe barrel (equipment) and a potential reaction of excipients in the waste receptacle. Six hazards, representing all four hazard categories (operational, environmental, equipment and human) were prioritized for immediate action (Table 6). Specific action recommendations for design interventions were made for risk mitigation and/or elimination.

Table 6. Selected Risks and Recommended Actions.

Hazard	**Category**	**Risk**	**Recommended Action**
Aggregation on injection	Operational	High	Inline washable filter, order formulation screen.
Multiple Peaks	Operational	High	Track multiple peaks, confirm peak id by DLS.

Plunger impact on injection	Operational	Serious	Physical pump stop to prevent damage, contact switch.
Reaction in waste receptacle	Environmental	Serious	Limit all screen components to MSDS level 2 or below.
Plunger barrel corrosion	Equipment	Medium	Implement cleaning procedure between experiments.
Burst fitting, sprayed formulation.	Personal	Medium	Require lab coats / goggles, pressure transducer.

Discussion

Although hazard analysis is well established in clinical biomedical engineering applications this case study provides an excellent opportunity to apply hazard analysis to the laboratory setting. During the hazard analysis, it became clear that both physical and functional (operational) hazard analysis should be equally considered in the research laboratory environment. The classical approach to hazard analysis focuses primarily on physical hazards, which have the potential to harm people, property and the environment. From a safety perspective, physical hazards are more important than operational hazards. However, in a laboratory environment, whether research or clinical, the accuracy of experimental data is necessary to successful operation. In research, accurate experimental results are required to establish the validity of the science being performed. In the clinical environment, accurate experimental results are critical to the proper diagnosis and treatment of patients. In this case, operational hazards can significantly overlap with physical hazards in terms of harm to people. Thus, operational or functional

validity is a necessary and important aspect of biomedical hazard analysis and risk assessment, regardless of the environment.

At the beginning of a biomedical engineering development project system validation is the primary interest. The question of whether or not a system will perform its designed function and whether or not it will improve upon an existing evaluation or treatment method is a fundamental question. Hazard analysis techniques directly address the validation question by requiring the researcher to ask, "What could go wrong with this system, and what is the likelihood and consequence of each potential failure." The conceptual and preliminary hazard analysis techniques establish an initial list of potential hazards and consequences. This list provides insight into the types of experiments and testing that should be performed in order to validate the system. For example, in the HSC hazard analysis complete replacement of one formulation with the next was identified as a requirement for accurate evaluation of protein-protein interactions in each formulation. Therefore flushing characteristics of the system were tested as part of the validation process. The components identified as critical to safety and function by various hazard analysis techniques will be candidates for verification as the project transitions from development to production and operation.

Regulatory and best practices standards provide important guidelines for developing a hazard analysis framework. Relevant regulatory agencies for the biomedical research laboratory environment in the United States include the Food and Drug Administration (FDA), Occupational Safety and Health Administration (OSHA) and the Environmental Protection Agency (EPA). It is necessary to review regulations promulgated by these agencies throughout the hazard analysis process. The regulations

most appropriate to our laboratory environment related to the chemical safety of formulation conditions. Regulatory requirements to control such hazards are generally personal protective equipment and barrier mechanisms, neither of which address the underlying cause of a hazard. Hazard analysis best practices, such as those described by ISO 14971[5] and ANSI [1], outline additional steps to identify the cause of hazards and eliminate the hazard through design changes throughout the development process.

Successful implementation of hazard analysis, in any setting, requires multiple stages of implementation. For this study the techniques implemented include preliminary hazard list (PHL), preliminary hazard analysis (PHA), Functional Hazard Analysis (FuHA), Failure Mode and Effect Analysis (FMEA). These techniques cover all four steps of hazard analysis – hazard identification, risk assessment, risk reduction and acceptance. There are alternative techniques for each stage of hazard analysis that may be employed in other biomedical applications. Here we discuss the benefits of each analysis technique chosen and its contribution to the hazard analysis.

First, the preliminary hazard list is an enumeration of potential hazards in the system based on the initial conceptual design. At this stage of development the design – and inputs to the PHL -- includes only rough outline of operations and concept sketches. This phase provides the best opportunity to make design changes with minimal impact on project cost. In the design concept phase there specific causes and failure modes are not identified. Instead, general hazard sources are identified based on the conceptual design.

Next, the preliminary hazard analysis (PHA) is used to assess hazard risks for the first time. In order to assess risk three factors must be determined: 1) hazard severity

and probability scales, 2) cause and effect for each hazard, 3) probability of each hazard to cause harm and severity of that harm. Hazard analysis / risk assessment standards[1,4,5] recommend logarithmic scales for severity and probability; however the range will depend on the system. Budget constraints for this study mandated a catastrophic equipment failure above $50,000. By contrast, a commercial application might only consider a catastrophic equipment failure to be one over $500,000. The probability and severity scales chosen in Tables 4 and 5 are based on the HSC system. It is important to note that probability and severity scales are dependent on the system to be analyzed.

The probability of each hazard to cause harm and the severity of that harm cannot be determined until a specific failure state is identified. After cause and effects of each failure state (whether physical or functional) are determined probability of the cause to occur and severity of the effects can be estimated. After the probability and severity scale is established and causes and effects are determined for each failure state the probability and severity for the hazard is assigned. During PHA the failure states are based on the preliminary design specification and expert opinion of the hazard analysis team. After each step of the design process accompanying hazards are review and recommended actions taken according to the priority established by the decision matrix (Table 7).

Table 7. Risk acceptance decisions were based upon a Decision Matrix.

Level of Action Required	Mishap Risk Indices	
	Equipment and Functional Risks	Human Injury Risks
High Risk *Immediate action needed*	1-3	1-5
Serious Risk	4-6	6-9

Action needed as soon as possible		
Moderate Risk *Action needed some time in the future*	7-13	10-13
Low Risk *No action needed*	14-16	14-16

At the system design level a functional hazard analysis (FuHA) identifies and evaluates functional hazards in the system. The information required for FuHA is: a list of all functions in the system, function flow and the preliminary hazard list and analysis [Ericson]. Each function in the list is evaluated for potential failures, incorrect behavior or invalid data. The FuHA is particularly well suited to identify hazards in software modules as they are primarily functional components -- for example the failure of the software to detect a protein elution peak.

The next design level hazard analysis technique applied to the HSC system is failure mode and effect analysis (FMEA). The FMEA is a detailed design hazard analysis performed at the component level where the component can mean individual hardware components, assemblies and subsystems or functions. While there is overlap between FuMA and FMEA, the latter is considered a bottom up approach. The approach is used here to focus on more detailed analysis of individual components of the HSC.

Additional hazard analysis techniques considered were Fault Tree Analysis (FTA) in the system design phase and Hazard and Operability Analysis (HAZOP) in the detailed design phase. In the design phase FuHA was chosen over FTA, a detailed failure event analysis. The functional approach was considered more applicable to the system because of the relatively low complexity of the failures identified in the preliminary hazard list

and preliminary hazard analysis. HAZOP, a technique used to analyze chemical processes in chemical engineering applications, was determined to be unnecessary in the detailed design phase. This was determined after the decision was made to use only chemicals with MSDS hazard category of two or below and to only use FDA approved inactive ingredients for pharmaceutical formulation screening. Inherently Safer Design (ISD) techniques focus on chemical safety by reducing storage quantities, safe storage, and reduction of potential interactions of chemicals and are appropriate for projects with more complex chemical safety concerns.

Actions taken for risk reduction include both procedural [6] and design [10] changes. An example of a design improvement to our system prompted by FMEA was pressure monitoring and disabling of pumps when system pressure rises above a set level. This improvement reduced the probability of a pressure burst and improved the risk index of the hazard from 9 to 13 by improving both a human risk (chemical exposure) and an operational expense (invalidation of the experiment). An example of a procedural improvement was the re-ordering of formulation conditions such that excipients, which more commonly cause protein precipitation, are tested last. An increase in the number of formulations tested before precipitation is observed decreases the operational impact associated with protein precipitation in the system and reduces the risk index of the hazard from 1 to 9. These examples demonstrate systematic improvements in the design and operating procedures in a biomedical research environment.

The flexibility of hazard analysis makes it customizable to different lab environments -- whether clinical or experimental -- throughout a device or process lifecycle. The preliminary hazard list and preliminary hazard analysis (PHL/PHA) is

used to identify major potential hazards of the primary system components. In the HSC system functional hazards were quickly identified to be the most important. In the design of an implantable device or when evaluating procedures in a clinical laboratory human safety hazards would be identified as higher priority due to increased probability and severity estimates.

Regardless of whether or not a product or process is being developed or evaluated in a laboratory or field setting the PHL/PHA provides the basis for further hazard analysis techniques. The selection of subsequent techniques depends on the identified hazards and the current design phase. The HSC design, being in the development stages, benefited from FuHA and FMEA. Functional hazard analysis organizes hazards from a top-down approach and is useful when evaluating multi-step laboratory and device processes. For the HSC system FuHA was used to identified potential hazards with the data analysis and software control systems during the conceptual and testing design phases. The FuHA was also helpful for developing standard operating procedures that will be used in the transition from development to production and operation. FMEA, on the other hand, focuses on a bottom-up approach examining individual components. FMEA is more applicable during the detailed design phase of development when components are specified and to establish maintenance schedules. Although this paper focuses on hazard analysis in the development process, an established laboratory environment can use hazard analysis for existing equipment and processes to improve the safety and effectiveness of operating procedures.

Conclusion

A structured hazard analysis and risk assessment process provides a systematic way to identify and prioritize the mitigation of risks to people, the environment, equipment and operations within a system. Here we have applied hazard analysis and risk assessment in a biomedical research laboratory environment to improve the safety and productivity of a high-throughput analysis system in the field of drug design and optimization. Early stage hazard analysis includes the first step of hazard tracking – identifying and recording hazards (Appendices 1-5). These improvements demonstrate that best practice standards such as ANSI/GEIA-0010-2009 are relevant and applicable to research environments.

This case study covers hazard analysis in the design stages of a project where changes to the design are able to improve safety and reliability with the least amount of cost. It is important to note, however, that hazard analysis is a continual process which must be revisited during iterative design changes and as additional information on failures in the field is collected. By collecting and maintaining information about hazards even when the risk has been reduced future changes and information can be evaluated in the context of expected hazards. In fact, the FDA requires that medical device failures are reported in a timely manner through the Manufacturer and User Facility Device Experience (MAUDE) database. Even in the case of research laboratories a basic failure tracking system combined with the linkage of expected failures to specific functions and components is expected to help identify potential system improvements.

Although the quantitative aspects of our hazard analysis, such as the quantitative hazard frequency and equipment severity scales are specific to our high-throughput system, the methods are expected to be applicable to both clinical and research laboratory environments. In this study standard hazard analysis techniques in the early design phase (PHL, PHA) identified many potential improvements in the system. Another common technique for detailed hazard analysis, FMEA, also identified specific component improvements. We found the most useful hazard analysis technique in the biomedical research laboratory to be the functional hazard analysis. The evaluation of each system function identified hazards which could cause harm to the proper functioning of the HSC system.

References

1. ANSI GEIA-STD-0010. Standard Best Practices for System Safety Program Development and Execution. October 2008. Tech America: Arlington, VA.

2. Bonfant, G. et al. Clinical risk analysis with failure mode and effect analysis (FMEA) model in a dialysis unit. *J. Nephrol* **23**, 111-118 (2010).

3. Chiozza, M.L. & Ponzetti, C. FMEA: A model for reducing medical errors. *Clinica Chimica Acta* **404**, 75-78 (2009).

4. Ericson, C.A. *Hazard Analysis Techniques for System Safety*. (Wiley-Interscience: 2005).

5. "Good Laboratory Practice Standards." Code of Federal Regulations Title 40, Part 792, 2009.

   ISO 14971:2007. Medical devices -- Application of risk management to medical devices. 2007.

6. Krouwer, J.S. An improved failure mode effects analysis for hospitals. *Arch. Pathol. Lab. Med* **128**, 663-667 (2004).

7. Manuele, F.A. *Advanced Safety Management Focusing on Z10 and Serious Injury Prevention*. (Wiley-Interscience: 2008).

8. NIOSH. System Safety and Risk Management – NIOSH Instructional Module. March 1998. Center for Disease Control: Cincinnati, OH.

9. Paté-Cornell, E. & Boykin, R. Probabilistic risk analysis and safety regulation in the chemical industry. *Journal of Hazardous Materials* **15**, 97 - 122 (1987).

10. Patel, S.M. et al. Methods of failure and reliability assessment for mechanical heart pumps. *Artif Organs* **29**, 15-25 (2005).

11. Rosenberg, A.S. Effects of protein aggregates: An immunologic perspective. *AAPS J* **8**, E501-E507 (2006).

12. "Toxic and Hazardous Substances." Code of Federal Regulations Title 29, Part 1910, Subpart Z, 2009.

13. Veterans Administration National Center for Patient Safety. The Basics of Healthcare Failure Mode and Effect Analysis. Course Notes. available at http://www4.va.gov/ncps/SafetyTopics/HFMEA/HFMEAIntro.doc. accessed April 23, 2010.

PROTEIN SOLUBILIZATION: A NOVEL APPROACH

DAVID H. JOHNSON, W. WILLIAM WILSON, LAWRENCE J. DELUCAS

Submitted to Journal of Chromatography B

Format adapted for dissertation

Abstract

Formulation development presents significant challenges with respect to protein therapeutics. One component of these challenges is to attain high protein solubility (> 50 mg/ml for immunoglobulins) with minimal aggregation. Protein-protein interactions contribute to aggregation and the integral sum of these interactions can be quantified by a thermodynamic parameter known as the osmotic second virial coefficient (B-value). The method presented here utilizes high-throughput measurement of B-values to identify the influence of additives on protein-protein interactions. The experiment design consists of three tiers of additive screens used to ultimately determine novel solution conditions that improve protein solubility. The first screen identifies individual additives that reduce protein interactions. A second set of B-values are then measured for different combinations of these additives via an incomplete factorial screen. Results from the incomplete factorial screen are used to train an artificial neural network (ANN). The "trained" ANN enables predictions of B-values for more than 4,000 formulations that include additive combinations not previously experimentally measured. Validation steps are incorporated throughout the screening process to ensure that 1) the protein's thermal and aggregation stability characteristics are not reduced and 2) the artificial neural network predictive model is accurate. The ability of this approach to reduce aggregation and increase solubility is demonstrated using an IgG protein supplied by Minerva Biotechnologies, Inc.

Keywords

Self-interaction chromatography, protein solubility, physical stability, high-throughput screening, neural network, design of experiments

Abbreviations

ANN Artificial Neural Network

BCA Bicinchoninic Acid

DLS Dynamic Light Scattering

Fab Antibody Fragment Region

FDA Food and Drug Administration

GLM General Linear Model

HSC High-throughput Self-interaction Chromatography

IgG Immunoglobulin G

Mab Monoclonal Antibody

MES 2-(*N*-morpholino) ethanesulfonic

PBS Phosphate Buffered Saline

R_h Hydrodynamic Radius

RMSE Root Mean Square Error

Introduction

Protein therapeutics is the fastest growing class of drugs in the pharmaceutical industry [1]. As basic research reveals biological pathways, protein deficiencies/abnormalities and key protein targets involved in disease states, we gain a greater understanding of how protein molecules can be used to influence those diseases. The first clinically demonstrated example of a protein therapeutic involved replacement of human insulin for diabetes treatment [2]. Even in those cases that do not involve direct protein replacement, engineered immunoglobulins or other proteins have been demonstrated to influence disease pathways [3].

To achieve an effective protein dose using a small injection volume, the concentration of clinically-approved protein drugs typically exceeds 100 mg/ml [4]. Physical stability of the protein at these high concentrations is a critical concern due to potential for an immune response to protein aggregates [5]. Therefore, solubility and physical stability are key variables that must be addressed when developing a protein therapeutic. The International Conference on Harmonization (ICH), whose objective is to harmonize requirements for safety, effectiveness and quality for pharmaceutical products, includes molecular characterization as one factor of a stability indicating profile [6]. To maintain physical stability at high concentration, a large number of different formulations must be evaluated followed by additional optimization of promising candidates – a process that may require several months, consuming significant quantities of protein. A process is needed to efficiently evaluate solubility behavior of protein drug formulations. The following presents a general method to improve colloidal stability of protein solutions based on a novel high-throughput HPLC system and a sequential design

of experiments (DOE) that provides rational screening of formulations (beginning with individual buffered additives and progressing to complex formulations).

Evaluation techniques used to determine protein solubility behavior can be grouped into two broad categories; 1) detection of aggregates and 2) measurement of protein-protein interactions. Examples of aggregate detection methods include size-exclusion chromatography, analytical ultracentrifugation and dynamic light scattering (DLS). The integral sum of protein-protein interactions (as opposed to site-specific interactions) is typically quantified by a thermodynamic parameter known as the osmotic second virial coefficient (also referred to as the B-value) [7–10]. The B parameter is correlated with both protein solubility [11–13] and the ability of a protein to crystallize [11,13–15,15–17] depending on whether the protein-protein interactions are repulsive (positive B values) or attractive (negative B values), respectively. Several methods exist to measure B values, including osmotic pressure [18], analytical ultracentrifugation [19], static light scattering [14,20] and self-interaction chromatography [8] (SIC).

We previously described the development process of a novel, high-throughput instrument that uses an established technique, self-interaction chromatography (SIC) to rapidly measure B-values [8,9,21,22]. SIC requires that the protein of interest is randomly bound to column media (static phase) while the column is equilibrated with the formulation being tested. A small bolus of protein (1 uL) is injected onto the column and its elution volume/time measured by UV absorption. The volume required to elute the protein of interest from the column is related to the protein-protein interactions of the injected bolus of protein with randomly oriented stationary protein. Other variables affecting B-value calculations such as the concentration of bound protein and nonspecific

protein-media interactions are discussed in the methods section. The initial protein-media binding step requires up to 2mg of purified protein with each experimental B-value measurement in different test formulations consuming an additional 1 ug of protein. The time required for each measurement is approximately two hours (~30-45 minutes for protein elution and an additional hour for re-equilibration in a new formulation condition). The advantages of SIC over other methods to calculate B-values include: 1) relatively straightforward, automatable experimental protocol that does not require specific operator expertise/experience, 2) low total protein consumption, 3) dead column reference and multiple B value measurements per run and 4) ability to use the technology for both aqueous and membrane proteins.

The high-throughput self-interaction chromatography instrument (HSC) described in the following sections consists of several key technological improvements including: 1) simultaneous use of four columns, 2) miniaturization of column dimensions resulting in a significant reduction in total protein consumption, 3) a high capacity formulation reservoir with automated robotic dispensing of different formulation conditions, 4) automated data acquisition. This lab previously reported the hazard analysis process which was used in the development of the HSC [23]. Use of the HSC instrument is combined with an approach that includes multiple levels of formulation screens and validation steps to assess: a) additive influence on protein thermal stability, b) protein-protein interactions and c) identify formulations with improved solubility.

Material and Methods

Multi Tiered Additive Screen

The approach described in this paper provides a rapid, cost effective method to determine solution conditions that optimize protein solubility. It requires 10 to 30 mg of protein and two to six weeks depending on the number of conditions needed to optimize solubility. The approach includes multiple levels of high throughput screening along with validation steps at each stage of the screening process. The overall goal of the approach is to determine which combination and concentration of additives minimize protein-protein attraction, thereby increasing protein solubility.

The flow chart presented in Figure 2 shows that the entire formulation screen is performed in five distinct phases: baseline measurements, initial screen, incomplete factorial screen, neural network modeling and confirmation with integrated quality control steps. Each screening step is described in more detail in the following sections.

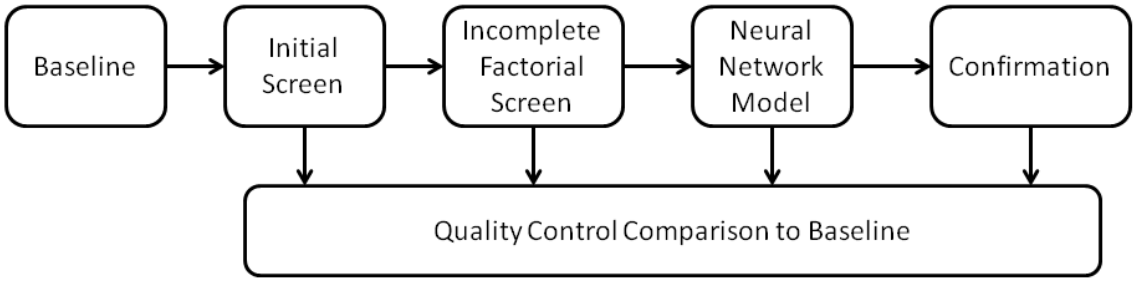

Figure 1. Experiment Flow. Flow chart of the experiments conducted to identify formulations with reduced aggregation. Most steps are followed by two actions: validation and comparison with baseline and the next experiment in the screening process.

In the first stage, additives are screened individually via experimentally measured B-values to identify those most effective at reducing protein-protein interactions. More

than 300 additives are approved by the FDA for use in injectable drug formulations as inactive ingredients [24]. If all 300 were to be assessed, just three concentration levels of two additives at three pH levels would result in over 1.2 million possible formulation conditions. To reduce the number of individual measurements needed to optimize a protein's solubility, our initial screen simply identifies which additives individually reduce protein-protein interactions. Those additives are then combined using an orthogonal array (assuring combinations of additives are equally represented throughout the screen) into formulations based on a balanced combination of each additive and additive concentration. For each protein studied a numerical model of how additives affect protein-protein interaction (B-value) is created by training an artificial neural network (ANN) using experimental data generated from the balanced screen. The neural network model is used to predict the B-value of the full factorial of screened additives. The following experimental results demonstrate the ability of an HSC system combined with this multi-tiered screening process to rapidly determine formulations with improved solubility behavior for a candidate IgG protein therapeutic compared to its original formulation.

High-throughput Self-interaction Chromatography (HSC)

HSC requires covalent attachment of the protein of interest (~2 mg) to media contained in a chromatography column. The same protein is injected into the mobile phase and retention times measured. Protein retention time measurement in each formulation requires approximately 1 ug of additional protein per column. Each

experiment requires approximately two hours using traditional column sizes and a precision HPLC. To reduce experiment time and protein consumption compared to traditional B measurement techniques, this laboratory reduced column dimensions (0.5mm x 180mm) and developed an automated high-throughput screening system to enable uninterrupted, simultaneous data acquisition from four chromatographic columns.

A schematic drawing of the HSC system is shown in Figure 1. Four valves control fluid flow throughout the system including the formulation, injection, and bypass valves. The formulation valve controls access of the formulation pump to the reservoir or other portions of the system. This pump withdraws formulations from the reservoir to the formulation syringe which then extrudes these solutions over one of the four columns. Bypass valves enable flushing the entire system with new formulations. This process is fully automated, accommodating 48 different formulations in 10 mL reservoirs. After column equilibration with new formulations, the protein is automatically injected over each column via the injection valve. The retention time of the eluted protein is determined using a UV280 detector.

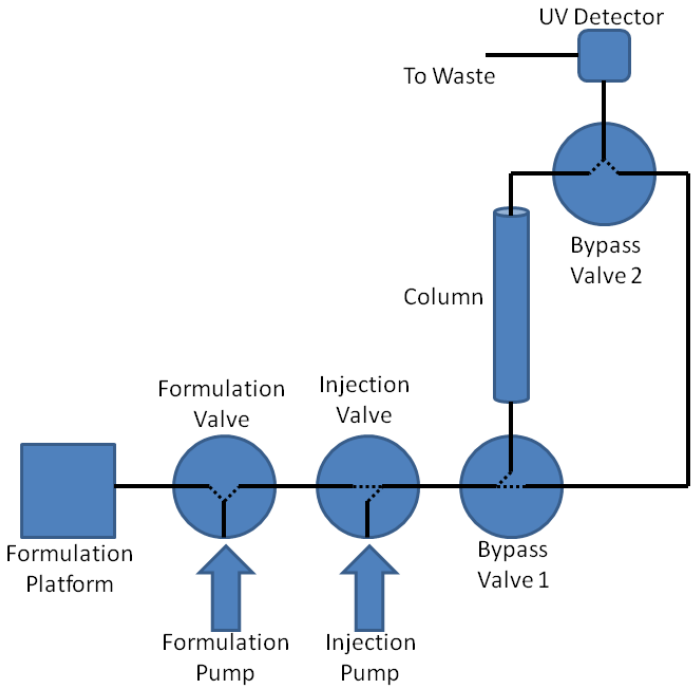

Figure 2. Schematic drawing of the HSC system. For clarity, this figure shows only a single flow path. The valves, platform and detectors handle four channels in parallel.

Baseline Measurements. Protein purity is initially confirmed via SDS PAGE and staining with BioSafe Coomassie. Protein B-value measurements are established using a "reference buffer" solution which provides a baseline reference for subsequent comparisons to identify formulation improvements. The reference buffer is typically a previously identified "good formulation" solution, judged by the extent of nonspecific protein aggregation using size exclusion chromatography (SEC) and/or dynamic light scattering (DLS). To ensure conformational consistency of the protein, the unfolding temperature is measured using a MicroCal VP-capillary differential scanning calorimetry system. The B-value for the protein in the reference buffer can be measured via SIC using either a Shimadzu HPLC or the custom HSC system. This B-value measurement is the primary metric by which additives and formulations are evaluated. Baseline

measurements are repeated between screens to ensure that the integrity of the column is not adversely affected due to changes in capacity (due to irreversible protein binding) and degradation of the media and/or media packing (caused by excessive back-pressure, non-specific irreversible protein binding). The reference buffer for the Fab protein presented in this paper is phosphate buffered saline (PBS -- diluted 1x contains 137 mM NaCl, 2.7 mM KCl, 10mM Na2HPO4, 2 mM KH2PO4, pH 7.4).

Self Interaction Chromatography

From a broad perspective, self-interaction chromatography is an affinity chromatography technique that involves binding the protein of interest to chromatography media, injecting a bolus of protein over the column, and measuring the retention time of the protein as it passes through and interacts with the protein bound to the media. Longer retention times are associated with increased protein-protein interaction [7,22]. The following sections provide a detailed description of the SIC method including media selection (based on evaluation of the concentration of bound protein), column preparation, retention time measurement and B-value calculation based on protein retention time.

Table 1. Chromatography Media Binding Test

	Binding Group	Binding pH	Reagents	Binding Time	Blocking	Blocking Time
Formyl	$-NH_2$	6.9 - 9.0	NaBH3CN	4h	Ethanolamine	1h
Tresyl	$-NH_2$, -SH	7.0 - 9.0	None	24h	Ethanolamine	4h
Amino	-HOOC, -OHC	4.5 - 6.0	EDC[1]/NHS[2]	24h	3HP[3]	4h
Carboxy	-NH	4.5 - 6.0	EDC[1]/NHS[2]	24h	Ethanolamine	4h

[1] ethyl(dimethylaminopropyl) carbodiimide
[2] n-hydroxysuccinimide
[3] 3-hydroxypropionic acid

Binding test. There are multiple protein binding chemistries available to covalently bind protein to media such that the point of attachment between the protein's surface and the media is random (this is accomplished by using free amine or carboxyl functional R-groups associated with amino acids). To identify optimal binding conditions, four different chemistries available from TosoHaas are considered: formyl, tresyl, amino and carboxy. For each media type, 20 ul of beads are washed three times with 1mL of the appropriate binding buffer for the given media (MES or phosphate buffer at pH 6.0 or 8.0 depending on the media). The binding tests are performed by adding 20 ul of 5 mg/ml protein to each sample of washed beads. Reagents are added according to the given chemistry. Table 1 displays the media options tested, their binding formulations, reagents required and binding times. After the binding process is complete, the media, now bound with some quantity of protein, is washed again three times with 200 uL of the binding buffer. A BCA assay is performed on each media sample to precisely quantify the amount of bound protein. This process requires ~0.4 mg of protein (20 ul * 5 mg/ml * 4 tests). In addition to identifying the media with the highest binding capacity, DSC is performed on the protein in the binding buffer, as a quality control step, to confirm that

the binding buffer does not destabilize or partially denature the protein. The binding chemistry with the highest quantity of bound protein, as determined by the BCA assay, is used for subsequent preparation of the SIC columns.

Binding and packing a column. For each protein investigated, the protein chemical binding process is scaled up 15-fold to support simultaneous use of three "live" columns ("live" column = column with protein). A fourth "dead" column ("dead" = column without protein) is prepared to serve as a control to assess protein interactions with the chromatographic media. The binding process for the live columns consists of adding 300 uL of 5 mg/ml protein to 300 uL of media. The appropriate reagents are added to the protein media solution followed by placement of samples on a rotating mixer for an appropriate amount of time according to Table 1. After completion of the protein binding step, the media is washed three times with 2 mL of the binding buffer. Using the identical binding reaction, remaining active sites of the media are capped with a capping reagent (0.5 M ethanolamine or 3-hydroxypropionic acid) used in place of the protein solution. At this point, 100uL of media is prepared with the capping reagent (no protein) to produce a capped-only "dead" column media. It should be noted that the covalently bound protein assures several different orientations on the column media due to the fact that the protein's covalent bond is formed with free amine nitrogens (which are typically available at multiple positions along the proteins exterior surface). Thus SIC-determined second virial coefficients are not less due to a specific protein-media orientation that would prevent interaction with various regions on the protein's surface.

The prepared media is packed into 0.02" i.d., 1/16" o.d. and 20 cm length columns consisting of teflon FEP tubing (IDEX) sealed at one end with a union (Valco

ZU1CFPK) containing a 2 micron frit (Valco 2SR1-10). Two 1cm sections are cut from the end of each column for protein analysis via Pierce BCA assay to determine the protein concentration bound to the column media. The concentration of bound protein is a critical variable required to calculate B from the measured protein retention time. Generally, for a 150 kDa protein (IgG) a minimum binding concentration of 5 mg/ml is required to yield a sufficient number of protein interactions to provide the sensitivity needed to accurately measure B-values for each solution condition. Finally the packed column is sealed with a union containing a 2 micron frit.

To separate injected protein from the storage buffer in which it is contained a desalting "guard" pre-column is also prepared for each live column and the dead column. This ensures the protein is fully equilibrated in the formulation of interest before it passes over the SIC column. The 60 cm, 0.03" i.d., 1/16" o.d. teflon FEP guard column uses Sephadex G-25 media. The media is prepared by soaking in a phosphate buffered saline (PBS) solution, pH 7.4, for 4 hours and rinsed 4 times with 5 mL of PBS. The media is packed into the column tubing and sealed at both ends as described for the SIC columns.

Retention time measurements. A guard column is positioned prior to each live and dead column to separate absorption peaks caused by salts, buffer and other small molecules contained in the base buffer used to solubilize the protein (these small molecules sometimes absorb at or close to 280nm and co-elute close to the protein absorption peak). The system is equilibrated in the formulation of interest by passing approximately 5 system volumes of formulation solution through the columns. After equilibration, a standard affinity chromatography experiment is performed by injecting the protein of interest in the mobile phase formulation at a constant 8 uL / min flow rate and protein

elution time is measured from the time of injection until the time of elution from each column (protein elution is detected via U.V. absorbance at 280 nm). The retention time is defined as the time required for the protein to pass through the SIC column only -- total elution time through both columns minus the elution time through the guard column, measured separately. This process is repeated for each of the additive formulations contained in the initial and incomplete factorial screens.

DSC Confirmation

B-values for protein in each formulation are measured followed by assessment of protein integrity for the nine additives yielding the most positive B-values (this is assessed using DSC measurements of protein thermal stability). This step serves as a quality control measure to ensure that protein denaturation is not a significant factor affecting protein retention times. Samples are prepared for DSC by buffer exchange using a centrifugal concentrator. A quantity of 0.25 mg protein from solution is added to the centrifuge tube. Three mL of formulation is added and spun down at 3000rpm in an Eppendorf-5810R centrifuge to 500 uL and this is repeated four times for a buffer exchange of over 99.9%. The final effluent is used as the blank control for the DSC experiment. If a formulation fails the DSC confirmation step (a drop in unfolding temperature of more than 4^{o}C) the next most positive B-value formulation is evaluated until nine additives are identified that improve B-value without significantly reducing protein unfolding temperature.

B Value Screens

Initial screen. The initial screen is used to identify individual additives that contribute to protein-protein repulsion (positive B-value). The formulations for the initial screen are prepared by combining 10x concentrated phosphate buffered saline (PBS -- diluted 1x contains 137 mM NaCl, 2.7 mM KCl, 10mM Na3PO4, 2 mM KH2PO4), and a single additive from concentrated stock, filled to 90% volume with Millipore H_2O and titrated with NaOH or HCl to pH 7.4. The list of forty additives, each tested individually in the initial screen, can be found in Appendix A. The B-value of the protein of interest is measured in each formulation using SIC. The additives most beneficial to reduce protein-protein interactions are not known a priori. Therefore, the purpose of the initial screen is to identify additives that reduce protein-protein interaction for a given protein. Thermal stability of the protein combined with those individual additives exhibiting the nine highest B-values is then confirmed using DSC. Three salts and six additives with the highest B-values which also maintain thermal stability via DSC (defined as those additives that do not result in more than a negative 4^oC temperature shift) are chosen for the incomplete factorial screen.

Incomplete factorial screen. After completion of the DSC tests, the top nine solutions are combined in an incomplete factorial using an orthogonal array according to the Taguchi method [25]. This method ensures that additive identity and additive concentration are equally represented throughout the incomplete factorial screen. The screen specifies more complex formulations with multiple additive components at high, medium and low concentrations. B-values of the protein in each formulation of the incomplete factorial screen can be measured by SIC using either the HSC system or a Shimadzu HPLC.

Neural network training. The measured B-values are used to create a numerical model of how specific additives in the formulation affect protein-protein interaction. An artificial neural network (ANN) model [26] is first trained five separate times using a different random set of 4/5 of the data with the remaining 1/5 of the data used for validation. The weights of the neural network are adjusted based on the training set using a standard back-propagation algorithm until the error of the validation set (not used to adjust weights) is no longer improved by incremental changes to the ANN weights. This is an established method to train and prevent over-fitting of an ANN and is described in Bishop's book, Neural Networks for Pattern Recognition [27]. In a previous publication [28], our lab compared this method to that of a standard general linear model (GLM) and found the ANN to exhibit reduced prediction error compared to the GLM.

Neural network prediction and confirmation. The trained ANN returns a B-value given a formulation where the output is based on previously measured B-values. After training is complete the neural network is presented with each formulation in the complete factorial combination of all parameters measured. This consists of over 4,000 possible formulations that combine one or two additives, a salt and a buffer. The ability of the ANN to predict B-values of novel formulations is evaluated by B-value measurement (by SIC) of several predicted formulations, chosen throughout the predicted B-value range. A root mean square error (RMSE) is determined for the model predictions compared to the actual B-values measured. Confirmation measurements serve two purposes: an evaluation of the model predictions and confirmation of predicted formulations with improved B-values.

Thermal stability confirmation. The predicted formulations with increased B-value measurements are experimentally validated followed by DSC confirmation of protein thermal stability in these new formulations.

Indirect Solubility Testing. Dynamic light scattering (DLS) is used to indirectly test the solution solubility characteristics of the protein by evaluation of the aggregation properties of the protein as protein concentration is increased. Due to a limited supply of protein (the proprietary Fab protein, provided by Minerva Biotechnologies, Inc., used for these experiments was of limited supply), DLS measurements were restricted to the most positive B-value formulation (by prediction and measurement) and the most negative B-value formulation from the initial screen. The rationale is to determine if more positive B-value predicted formulations produce significant improvements in solubility behavior. Each protein formulation is filtered through a 0.22 micron filter into a 10 uL quartz 90^{o} scattering cell. Scattering counts are measured using a Wyatt DynaPro Titan DLS system to confirm that there is not a significant scattered light signal from the buffer itself. The protein is buffer-exchanged into each formulation using centrifugal concentrators. Diffusion coefficients and apparent hydrodynamic radii of the protein are measured in the solution at a fixed concentration. The Fab sample is concentrated with DLS measurements taken at various points during in the concentration process. The aggregation behavior of the protein is compared as a function of protein concentration for each formulation. The minimal use of protein required for DLS measurements and the fact that protein can be recovered for additional tests resulted in this being the method of choice for comparing aggregation behavior of the protein.

Calculation of B Value

The standard method of B value calculation, published by Tessier, Lenhoff and Sandler [8], is used in this research:

$$B = \frac{N_A}{MW^2}\left(V_{HS} - \frac{k'}{\varphi\rho}\right)$$

In this equation, N_A is Avogadro's number (molecules/mol) and MW is molecular weight (g/mol). V_{HS} (ml/molecule) is the hard sphere volume of the protein of interest. The parameter, φ (cm^2/mL), is the phase ratio of the media defined as the ratio of the surface area available to that of the volume available to a mobile phase protein passing through the media. Phi is a characteristic of the media used and molecular weight of the protein and has been determined for several media types including TosoHaas affinity media [29]. The parameter ρ (molecules/mL)is the amount of protein per unit volume bound to the media (determined via the BCA assay as described previously).

The final parameter k' is the retention factor. While the other parameters are fixed for a given column, the K' measurement is the primary variable associated with changes in protein-protein interactions. K' is calculated according to the following [8]:

$$k' = \frac{(V - V_0)}{V_0}$$

V is the retention volume of the protein over the live column and V_0 is the retention volume of an equivalently sized non-interacting marker. Previously an acetone marker has been used as the non-interacting marker and a correction was used to adjust for the small size of the acetone marker compared to the protein [8,9,30]. However, the addition

of the guard column used to equilibrate the protein in the mobile phase substantially shifts the acetone marker with respect to the protein marker. In the multi-column system, the dead column is used to identify the non-interacting retention volume. Therefore, V_0 is the retention volume of protein eluted from the dead column in the formulation of interest. This method has the added benefit of accounting for changes in protein-media interactions in the presence of different formulations.

Results

Fab protein was subjected to the three-tiered screening process. The specific antigen is proprietary and not known to our laboratory.

Minerva Fab – Initial Screen

The primary goal of the Minerva project was to produce a highly soluble storage solution for the protein, not a solution for direct injection. Therefore, this screen contains additives and concentrations not approved for human use. Identifying such a formulation is useful for long-term storage of protein produced on a large scale and for preliminary formulations that can be used with other structure-function analysis methods as well as for *in-vivo* animal studies. With different initial screen components the screening methodology and high-throughput technology are applicable to preparation of solution conditions for pre-clinical evaluation. Minerva provided our lab with ~25 mg of the Fab portion of a proprietary monoclonal antibody (Mab) being considered for future clinical

trials. It was assumed that if improved solubility conditions could be discovered for the Fab, these conditions would also exhibit improved solubility for the complete monoclonal antibody. Components and concentrations of the additives used in this screen can be found in Appendix A.

Table 2. Most positive B-values of Minerva Fab Initial Screen

Additive Name	B-value[1] Measured	Delta T_M (^oC)
Default Buffer - PBS	0.2	+0.0
1. 0.1M Arginine	5.8	+0.6
2. 400mM LiCl	5.1	-1.6
3. 400mM Na Thiocyanate	4.4	-3.7
4. 0.1M Arg., 0.1M Glu. Acid	3.2	NA
5. 400mM LiSO4	1.2	-9.0
6. 0.1M Glucose	0.8	-0.6
7. 400mM Na Citrate	0.1	+2.5
8. 0.1M Trehalose	0.0	-0.1
9. 10% (w/v) 1,6 Hexanediol	0.0	-10.5
10. 400mM NH_4 Citrate	-0.6	-2.3

[1] (* 10^{-4} mol*ml/g^2)

Table 2 shows the additives producing the nine highest B-values chosen from the initial screen. This includes the additives that failed DSC confirmation (1,6-hexanediol and Li_2SO_4). These two additives were replaced with those producing the next most positive B-values, NaCl and Glutamic Acid. The additives chosen from the initial screen are applied to an orthogonal array [31] to determine the additives and concentrations used for each formulation condition in the incomplete factorial screen. A full list of the 36 formulations in this phase of the screen can be found in Appendix B and the most positive B-values identified in the screen are in Table 3.

Table 3. Most positive B-values From Minerva Fab Incomplete Factorial Screen

Formulation	B-value[1] Measured
Default Buffer – PBS	0.2
1. 0.1M Phosphate pH 7.4, 0.15 M NaCl, 0.2 M Na Citrate, 0.6 M NH_4 Citrate	5.6
2. 0.1M Phosphate pH 7.4, 0.15 M NaCl, 0.1 M Trehalose, 0.05 M Glucose	4.1
3. 0.1M Tris pH 8.0, 0.05 M LiCl, 0.4 M Na Citrate	2.7
4. 0.1M Phosphate pH 7.4, 0.05 M NaSCN, 0.4 M NH_4 Citrate, 0.15 M Arginine	2.1
5. 0.1M MES pH 6.1, 0.1 M NaCl, 0.05 M Glutamic Acid, 0.15 M Trehalose	1.9
6. 0.1M Phosphate pH 7.4, 0.05M LiCl, 0.15 M Trehalose	1.4
7. 0.1M Tris pH 8.0, 0.1 M NaCl, 0.1 M Trehalose, 0.15 M Arginine	1.4
8. 0.1M MES pH 6.1, 0.15 M NaSCN, 0.1 M Glutamic Acid	1.4

[1] (* 10^{-4} mol*ml/g^2)

From the 27 different neural networks trained, the 5 x 2 topography provides the smallest validation error across all validation sets for the Minerva Fab protein. The average validation error is 1.2 B units. The trained neural network produces a range of B-value predictions from -5.4 to 4.3 B units and 4 formulations from the top quartile of B-values are chosen to yield improved formulations. Different topologies represent a different number of variables considered for influence on B-value. It is expected that some topologies (those that consider too few or too many variables) would produce lower validation errors than others. The evaluation of multiple topologies is automated and does not require additional effort and accounts for the fact that the number of variables which influence B-value are expected to differ from protein to protein. The measured confirmation of B-value by SIC and change in unfolding temperature by DSC are given in Table 4.

Table 4. B-value Confirmations and DSC Unfolding Temperatures for Fab

Formulation	B-value[1] Predicted	B-value[1] Measured	Delta Tm (^{o}C)
1. 100mM Tris pH 8.0, 100mM NaCl, 150mM Trehalose, 400mM NH_4 Citrate	3.4	0.8	+5.4
2. 100mM Tris pH 8.0, 150mM NaSCN, 100mM Trehalose 600mM NaCitrate	1.6	1.6	+5.0
3. 100mM Tris pH 8.0, 150mM LiCl, 150mM Arg-HCl, 50mM Glucose	1.0	0.9	+0.2
4. 100mM Phosphate pH 7.5, 150mM NaCl, 100mM Trehalose, 100mM Glutamic Acid	4.1	2.1	+1.9

[1] (* 10^{-4} mol*ml/g^2)

The restriction on protein quantity received (25mg) limits the maximum solubility that can be determined for a given formulation. In the case of the Minerva Fab the formulations submitted to the company were tested by the company with larger protein quantities. Minerva concentrated the complete monoclonal antibody (Mab) in each formulation until visible precipitation was observed. These results are shown in Figure 3.

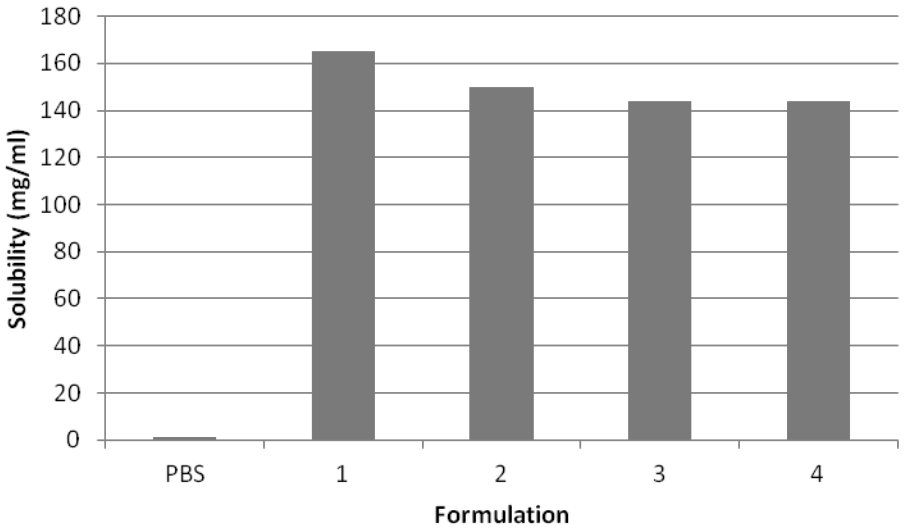

Figure 3. Solubility estimates of Fab from Minerva

Discussion

Comparison of results at each step in the screening process demonstrates the importance of each step in evaluating formulations. The following discussion compares the results of the two proteins evaluated for each step. Each of the following subsections is focused on a single step in the screening process outlined in Figure 1.

Baseline

Baseline measurements are important for both quality assurance (of the initial quality of the protein) and quality control (of formulation improvements). The baseline unfolding temperature provides a reference to quantify shift in unfolding temperature for protein equilibrated in each formulation. In the case of denatured protein, DSC does not result in a positive heat capacity signal and can be used to identify formulations which denature the protein. An additional quality assurance step includes evaluation of protein binding to a small quantity of each media type to identify the optimum binding chemistry for a particular protein. Both proteins bound to Tresyl media at greater than 7 mg/ml and demonstrated typical IgG unfolding temperatures from 60^{o}C to 90^{o}C. This ensures that the best media is chosen for protein binding.

Regarding use of PBS as the baseline formulation, it should be noted that antibodies are generally not optimally soluble in PBS, and that often a low pH buffer solution can improve solubility. This was not the case for the Fab obtained from Minerva. This protein and most other proteins evaluated are generally "problem" proteins (proteins that exhibit low solubility). The starting point for the formulation

screen for Minerva's Fab was the best available based on information provided by Minerva (the company approached Soluble Therapeutics due to their prior difficulty improving the solubility of the protein). Obviously, no single screening process can *always* produce and guarantee the maximally optimal formulation. That would require a complete factorial testing of the search space. As noted in the introduction, even with severe limitations on the dimensionality of the search space (2 additives, 3 concentration and pH levels) an exhaustive search of the space is not feasible. Soluble Therapeutics' screening process is designed to improve formulations with an approach that begins with identification of individual additives that influence protein-protein interactions followed by expansion to more complex formulations that contain multiple additives and additive concentrations. After individual additives are selected, an incomplete factorial additive screen is used to improve the chance of identifying combinations of additives that work well together.

Initial Screen

The initial screen of individual additives includes additives not in the FDA database. A wide range of additives was chosen in order to evaluate the overall system's capability to predict B-values that modify solubility behavior. Each additive formulation is screened using SIC to identify the B-value associated with the additive. Individual additives with the highest B-values from the initial screen are identified in Tables 2 and 5 for the Minerva Fab. Formulation optimization experiments performed with proprietary

proteins (both IgGs and other protein classes), using FDA-approved additives, resulted in alternative additives and additive concentrations.

As noted earlier, PBS was both the starting point for the initial screen and the "best" formulation available (defined as the base buffer) at the time. This is usually not the case. However, when a more optimally soluble formulation is available it is not necessarily the best starting point for the initial screen. Beginning with a complex formulation augmented with additives is likely to result in a search space located around a local maximum (but not necessarily the global maximum). Identification of a minimal buffer (such as PBS) allows for the inclusion of individual additives that may have a significant impact on protein-protein interactions; this may not be distinguishable in an already complex formulation. Thus, the minimal buffer formulation enables broadening of the search space to include more diverse formulations for the incomplete factorial screen.

Incomplete Factorial

The incomplete factorial screen design is based on experimental results from the initial screen. Additives identified in the initial screen (a different set for each protein) are combined using an orthogonal array to ensure equal representation of each additive throughout the screen. Each formulation in the incomplete factorial is evaluated by SIC to determine the B-value of the protein in the formulation. Appendices 3 and 4 identify all additives in the incomplete factorial for the Minerva Fab. The formulations from the incomplete factorial with the most positive B-values are listed in Table 3.

Neural Network

Neural network training produced a Fab validation error 1.2 B units and a percent validation error of 12.4%. After training, prediction of the complete factorial of additive combinations resulted in both positive and negative B-value predictions. The range of B-values predicted for the Minerva protein was -5.4 to +4.3 B units. Even before empirical confirmation, the broad range of these B-values included many formulations that were expected to have high solubility. Literature on the empirical relationship between B-values and solubility [11–13,32], indicates that protein solubility increases as B-value approaches zero. The solubility reported by Minerva (Figure 3) confirms that the formulations with high positive predicted B value have improved solubility.

Conclusion

The results demonstrate a significant improvement (i.e. more positive B-values) for the predicted formulations versus the base buffer as did Minerva's non-quantitative solubility assessment of the top four predicted formulations. The approach utilizes our novel technology (HSC) and design of experiments to evaluate multiple tiers of additive formulations. The results from these screens are evaluated with an artificial neural network model to identify formulations with improved solubility behavior. The formulations identified for the Minerva Fab improved solubility one-hundred fold over the existing baseline formulation and enabled the protein to advance to animal trials using two of the best predicted formulations.

It is important to emphasize the baseline checks throughout the process. Differential scanning calorimetry is essential to exclude false positives. After the initial screen of formulation additives, it is common to have at least one or two additives which reduce the unfolding temperature of the protein or result in denaturation, but result in positive B-values due to size exclusion effects. One could argue that a reduction in unfolding temperature of 2-3^{o}C is not significant in the thermal stability of a protein—especially if those temperatures are around 80^{o}C. However, a strong relationship between unfolding temperature and protein activity has not yet been evaluated for a significant number of proteins. In this screening process a cautious approach is taken to eliminate additives which could alter activity due to a conformational change. This is acceptable due to the large number of additives evaluated and the small incremental cost due to additional protein consumption. As more evidence is gathered regarding the nature of unfolding temperatures, solubility and activity, operating procedures will be adjusted. The evaluation presented in this paper provides a snapshot of a current screening technique to improve solubility behavior of therapeutic proteins. An aspect of this research that must be emphasized is that of solubility behavior vs absolute solubility. Absolute maximum solubility is difficult to measure for a protein solution because of the ability of a protein to super-saturate and result in significantly different solubility maximums in slightly different formulation conditions. Therefore, one approach that may prove useful is to evaluate the tendency of the protein to aggregate at increasing concentrations. Laser light scattering (and an observed A_{280} signal decrease) can be used to assess the tendency of different formulations to produce increasing protein aggregation. This technique allows protein solubility assessment with a small amount of

protein. To our knowledge, a method does not exist (when only small quantities of purified protein are available) to evaluate absolute protein solubility for proteins exhibiting a high maximum solubility.

The improvement of Mab solubility based on Fab screening suggests that problematic protein domains could be formulated separately to improve solubility of the complete protein. It is logical to assume that solubility improvements for a portion of a protein (i.e. protein domain that exhibits poor solubility) may also improve the solubility of the complete protein. Improvements to each step in the evaluation process can potentially reduce protein use and time. In addition to efficiency (time and protein) optimizations, we are exploring expansion of applications. For example an additional improvement would include use of circular dichroism as a reference to ensure that the protein structure is not significantly changing. This would be beneficial for therapeutic proteins at the pre-clinical stage in which end point results are undergoing additional analytical evaluations (e.g. the Minerva Fab).

Acknowledgements

The research being reported here was supported by NIH R01GM095639, NIH National Cancer Institute 5P30CA13148-40 and UAB Health Services Foundation (HSF). Therapeutic proteins for the study were provided by Minerva Biotechnologies, Inc. The authors Dr. William W. Wilson and Dr. Lawrence J. DeLucas of this publication have equity ownership in, serves as an advisor for, and serve on the board of director of directors for Soluble Therapeutics, Inc. which is developing products related to the

research being reported. In addition, the authors serve as a consultant to Soluble Therapeutics and receive compensation for Board of Director services. The terms of this arrangement have been reviewed and approved by the University of Alabama at Birmingham and Mississippi State University in accordance with its policy on conflict of interest in research.

Appendices

Supporting information includes two appendices containing the full list of formulations screened by self-interaction chromatography. They are the initial screen additives (Appendix A) and the incomplete factorial screen for Minerva Fab (Appendix B).

References

[1] P.V. Arnum, Evaluating the Pieces in Big Pharma's Manufacturing Investments, Pharmaceutical Technology. 37 (2013) 38–41.

[2] I.S. Johnson, Human Insulin from Recombinant DNA Technology, Science. 219 (1983) 632–637.

[3] B. Leader, Q.J. Baca, D.E. Golan, Protein therapeutics: a summary and pharmacological classification, Nat Rev Drug Discov. 7 (2008) 21–39.

[4] S.J. Shire, Z. Shahrokh, J. Liu, Challenges in the development of high protein concentration formulations, Journal of Pharmaceutical Sciences. 93 (2004) 1390–1402.

[5] H. Schellekens, Factors influencing the immunogenicity of therapeutic proteins, Nephrology Dialysis Transplantation. 20 (2005) vi3–vi9.

[6] International Conference on Harmonization, Quality of Biotechnological Product: Stability Testing of Biotechnological/Biological Products Q5C, (1995).

[7] B.L. Neal, D. Asthagiri, A.M. Lenhoff, Molecular Origins of Osmotic Second Virial Coefficients of Proteins, Biophysical Journal. 75 (1998) 2469–2477.

[8] P.M. Tessier, A.M. Lenhoff, S.I. Sandler, Rapid Measurement of Protein Osmotic Second Virial Coefficients by Self-Interaction Chromatography, Biophys. J. 82 (2002) 1620–1631.

[9] J.J. Valente, K.S. Verma, M.C. Manning, W.W. Wilson, C.S. Henry, Second Virial Coefficient Studies of Cosolvent-Induced Protein Self-Interaction, Biophys. J. 89 (2005) 4211–4218.

[10] W.G. McMillan, J.E. Mayer, The Statistical Thermodynamics of Multicomponent Systems, The Journal of Chemical Physics. 13 (1945) 276–305.

[11] C. Haas, J. Drenth, W.W. Wilson, Relation between the Solubility of Proteins in Aqueous Solutions and the Second Virial Coefficient of the Solution, J. Phys. Chem. B. 103 (1999) 2808–2811.

[12] S. Ruppert, S.I. Sandler, A.M. Lenhoff, Correlation between the Osmotic Second Virial Coefficient and the Solubility of Proteins, Biotechnol. Prog. 17 (2001) 182–187.

[13] B. Guo, S. Kao, H. McDonald, A. Asanov, L.L. Combs, W. William Wilson, Correlation of second virial coefficients and solubilities useful in protein crystal growth, Journal of Crystal Growth. 196 (1999) 424–433.

[14] A. George, W.W. Wilson, Predicting protein crystallization from a dilute solution property, Acta Crystallogr D Biol Crystallogr. 50 (1994) 361–365.

[15] F. Bonneté, S. Finet, A. Tardieu, Second virial coefficient: variations with lysozyme crystallization conditions, Journal of Crystal Growth. 196 (1999) 403–414.

[16] P.M. Tessier, H.R. Johnson, R. Pazhianur, B.W. Berger, J.L. Prentice, B.J. Bahnson, et al., Predictive crystallization of ribonuclease A via rapid screening of osmotic second virial coefficients, Proteins: Structure, Function, and Bioinformatics. 50 (2003) 303–311.

[17] P.M. Tessier, A.M. Lenhoff, Measurements of protein self-association as a guide to crystallization, Current Opinion in Biotechnology. 14 (2003) 512–516.

[18] Y.. Moon, C.. Anderson, H.. Blanch, J.. Prausnitz, Osmotic pressures and second virial coefficients for aqueous saline solutions of lysozyme, Fluid Phase Equilibria. 168 (2000) 229–239.

[19] S.A. Berkowitz, Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals, AAPS J. 8 (2006) E590–E605.

[20] P. Kratochvl, Classical Light Scattering from Polymer Solutions, Elsevier Science Ltd, 1987.

[21] T. Ahamed, M. Ottens, G.W.K. van Dedem, L.A.M. van der Wielen, Design of self-interaction chromatography as an analytical tool for predicting protein phase behavior, Journal of Chromatography A. 1089 (2005) 111–124.

[22] S.Y. Patro, T.M. Przybycien, Self-interaction chromatography: a tool for the study of protein-protein interactions in bioprocessing environments, Biotechnol. Bioeng. 52 (1996) 193–203.

[23] D. Johnson, M. Bidez, L. DeLucas, Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulation Equipment, Annals of Biomedical Engineering. 40 (2012) 898–906.

[24] Inactive Ingredient Search for Approved Drug Products, (n.d.).

[25] G. Taguchi, S. Konishi, Taguchi Methods Orthogonal Arrays and Linear Graphs: Tools for Quality Engineering, illustrated edition, Amer Supplier Inst, 1987.

[26] P. Marrone, Java Object Oriented Neural Engine, SourceForge. (n.d.).

[27] C.M. Bishop, Neural networks for pattern recognition, Oxford University Press, 1995.

[28] D. Johnson, A. Parupudi, W. Wilson, L. DeLucas, High-Throughput Self-Interaction Chromatography: Applications in Protein Formulation Prediction, Pharmaceutical Research. 26 (2009) 296–305.

[29] P. DePhillips, A.M. Lenhoff, Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography, Journal of Chromatography A. 883 (2000) 39–54.

[30] P.M. Tessier, S.D. Vandrey, B.W. Berger, R. Pazhianur, S.I. Sandler, A.M. Lenhoff, Self-interaction chromatography: a novel screening method for rational protein crystallization, Acta Crystallogr D Biol Crystallogr. 58 (2002) 1531–1535.

[31] A.S. Hedayat, N.J.A. Sloane, J. Stufken, Orthogonal arrays: theory and applications, Springer, 1999.

[32] K. Demoruelle, B. Guo, S. Kao, H.M. McDonald, D.B. Nikic, S.C. Holman, et al., Correlation between the osmotic second virial coefficient and solubility for equine serum albumin and ovalbumin, Acta Crystallographica Section D Biological Crystallography. 58 (2002) 1544–1548.

CONCLUSIONS

The manuscripts presented here demonstrate a consistent progression in the development of the HSC system. The first manuscript, subtitled, "Applications in Protein Formulation Prediction" uses a traditional HPLC instrument to measure B values. The unique contribution of the manuscript is to demonstrate the ability to train an artificial neural network to predict B values of novel formulations based on B value measurements. The limitations of this initial work are two-fold: 1) lysozyme is a relatively low molecular weight and inherently stable protein and 2) the low-throughput HPLC is only able to measure a single B value every two hours. The next two manuscripts address these issues through development of high-throughput hardware and evaluation of an antibody fragment.

The second manuscript, "Hazard Analysis and Risk Assessment in the Development of Biomedical Drug Formulations", applies hazard analysis techniques to the HSC system design. Traditional hazard analysis in the clinical environment focuses on patient health and safety. In development of the risk assessment procedure, applied to academic research, it became clear that experiment accuracy and integrity is the central concern. This process led to the introduction of a pressure sensor to identify problems with precipitating protein and the inclusion of a software specified cutoff when system pressure reaches a level that might damage the chromatography columns.

The final manuscript, “Protein Solubilization: A Novel Approach”, describes a comprehensive screening method that incorporates two additional screening tiers, a pre-screen to identify the chromatography media with the highest concentration binding and an initial screen to evaluate individual additives. The pre-screen is a quality control step used to evaluate the efficiency of four different protein binding chemistries that couple protein to the chromatographic media. The goal of this pre-screen is to find the optimum binding chemistry (that which results in the highest protein concentration bound to the media). The initial screen, which measures B values of individual additives in a minimal formulation, allows for evaluation of a greater number of unique additives than could be accommodated in the more complex incomplete factorial screen. Individual additives with the highest B values are chosen for the incomplete factorial screen. As in “Applications in Protein Formulation Prediction”, B values are measured for each formulation in the incomplete factorial and these results are used to train the artificial neural network. Predictions made by the trained neural network model for the complete factorial of combinations and a subset of the predictions are measured by self-interaction chromatography.

Hardware improvements that have contributed to the successful implementation of the HSC System include: dual-detectors, robotic formulation delivery and pressure monitoring. Placement of a detector before and after the SIC column (dual detectors) prevents guard column variations from affecting B value calculation. The robotic formulation delivery allows for uninterrupted evaluation of 48 formulations and pressure monitoring is able to identify over-pressure issues due to aggregation or other line obstructions.

Future Work

There are many potential topics of research both in system design and application. Primary areas of improvement in system design include screening methods, analysis and hardware. For screening methods, novel additives and sets of additives can be identified for particular protein or class of proteins, such as IgG. Identification of additive sets will require storage of formulation B values across different formulations and different proteins along with their protein class. A database to handle this information is currently under development. New hardware/method improvements include reduction in column diameter, monolithic columns (in-column polymerization) and on-column binding systems (binding protein to a pre-packed column) for a more simple experiment setup. On-column binding has been evaluated previously[23] for larger volume columns and we are in the process of evaluating on-column binding with the smaller diameter columns used in the HSC System.

The applications of the HSC technology include solubilization of viral capsids for vaccines and cross-interaction chromatography. One issue with evaluating B values of whole viral capsids is that they are typically at very low concentrations (micromolar). Based on the improved solubility of the Minerva IgG protein by solubilizing a sub-domain we are interested identifying whether or not constituent capsid proteins can be solubilized to improve whole capsid solubility. With cross-interaction chromatography instead of self-interactions, the interactions between two different protein molecules are evaluated. Identification of formulations that form stable protein complexes has potential use in both crystallization and storage stability. Overall, the multiple improvements made in the HSC system have resulted in higher throughput and accuracy while lessening

the need to manual intervention/monitoring. These accomplishments enable the experimenter to more efficiently explore the multiple applications of this technology.

REFERENCES

1. Johnson, I. S. Human Insulin from Recombinant DNA Technology. *Science* **219,** 632–637 (1983).

2. Arnum, P. V. Evaluating the Pieces in Big Pharma's Manufacturing Investments. *Pharm. Technol.* **37,** 38–41 (2013).

3. Shire, S. J., Shahrokh, Z. & Liu, J. Challenges in the development of high protein concentration formulations. *J. Pharm. Sci.* **93,** 1390–1402 (2004).

4. De Groot, A. S. & Scott, D. W. Immunogenicity of protein therapeutics. *Trends Immunol.* **28,** 482–490 (2007).

5. Schellekens, H. Bioequivalence and the immunogenicity of biopharmaceuticals. *Nat Rev Drug Discov* **1,** 457–462 (2002).

6. Schellekens, H. Factors influencing the immunogenicity of therapeutic proteins. *Nephrol. Dial. Transplant.* **20,** vi3–vi9 (2005).

7. Frokjaer, S. & Otzen, D. E. Protein drug stability: a formulation challenge. *Nat. Rev. Drug Discov.* **4,** 298–306 (2005).

8. Shire, S. J. Formulation and manufacturability of biologics. *Curr. Opin. Biotechnol.* **20,** 708–714 (2009).

9. Chirino, A. J., Ary, M. L. & Marshall, S. A. Minimizing the immunogenicity of protein therapeutics. *Drug Discov. Today* **9,** 82–90 (2004).

10. Center for Drug Evaluation and Research. Drug Approvals and Databases - Inactive Ingredients Database Download. at <http://www.fda.gov/Drugs/InformationOnDrugs/ucm113978.htm>

11. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. *Pharm. Res.* **20,** 1325–1336 (2003).

12. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W. & Katayama, D. S. Stability of Protein Pharmaceuticals: An Update. *Pharm. Res.* **27,** 544–575 (2010).

13. Ye, H. Simultaneous determination of protein aggregation, degradation, and absolute molecular weight by size exclusion chromatography–multiangle laser light scattering. *Anal. Biochem.* **356,** 76–85 (2006).

14. Demeule, B., Messick, S., Shire, S. J. & Liu, J. Characterization of Particles in Protein Solutions: Reaching the Limits of Current Technologies. *AAPS J.* **12,** 708–715 (2010).

15. Rhyner, M. N. The Coulter Principle for Analysis of Subvisible Particles in Protein Formulations. *AAPS J.* **13,** 54–58 (2010).

16. Neal, B. L., Asthagiri, D. & Lenhoff, A. M. Molecular Origins of Osmotic Second Virial Coefficients of Proteins. *Biophys. J.* **75,** 2469–2477 (1998).

17. Berkowitz, S. A. Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals. *AAPS J.* **8,** E590–E605 (2006).

18. Moon, Y. ., Anderson, C. ., Blanch, H. . & Prausnitz, J. . Osmotic pressures and second virial coefficients for aqueous saline solutions of lysozyme. *Fluid Phase Equilibria* **168,** 229–239 (2000).

19. Tessier, P. M., Lenhoff, A. M. & Sandler, S. I. Rapid Measurement of Protein Osmotic Second Virial Coefficients by Self-Interaction Chromatography. *Biophys J* **82,** 1620–1631 (2002).

20. Bakshi, M. & Singh, S. Development of validated stability-indicating assay methods--critical review. *J. Pharm. Biomed. Anal.* **28,** 1011–1040 (2002).

21. McMillan, W. G. & Mayer, J. E. The Statistical Thermodynamics of Multicomponent Systems. *J. Chem. Phys.* **13,** 276–305 (1945).

22. Zimm, B. H. Application of the Methods of Molecular Distribution to Solutions of Large Molecules. *J. Chem. Phys.* **14,** 164–179 (1946).

23. Rakel, N., Schleining, K., Dismer, F. & Hubbuch, J. Self-interaction chromatography in pre-packed columns: A critical evaluation of self-interaction chromatography methodology to determine the second virial coefficient. *J. Chromatogr. A* **1293,** 75–84 (2013).