Journal of Pharmaceutical Sciences xxx (2016) 1–9

Contents lists available at ScienceDirect

Journal of Pharmaceutical Sciences

journal homepage: www.jpharmsci.org

Pharmaceutics, Drug Delivery and Pharmaceutical Technology

Quality by Design Approaches to Formulation Robustness—An Antibody Case Study

Christine Wurth [1,*], Barthelemy Demeule [2], Hanns-Christian Mahler [1], Michael Adler [1]

[1] Pharmaceutical Development & Supplies, PTD Biologics Europe, F. Hoffmann-La Roche Ltd., Basel, Switzerland
[2] Pharmaceutical & Process Development, Genentech Inc., South San Francisco, California 94080

ARTICLE INFO

Article history:
Received 17 November 2015
Revised 1 February 2016
Accepted 16 February 2016

Keywords:
biotechnology
formulation
HPLC
multivariate analysis
oxidation
factorial design
protein aggregation
protein formulation
proteins
stability

ABSTRACT

The International Conference on Harmonization Q8 (R2) includes a requirement that "Critical formulation attributes and process parameters are generally identified through an assessment of the extent to which their variation can impact the quality of the drug product," that is, the need to assess the robustness of a formulation. In this article, a quality-by-design–based definition of a "robust formulation" for a biopharmaceutical product is proposed and illustrated with a case study. A multivariate formulation robustness study was performed for a selected formulation of a monoclonal antibody to demonstrate acceptable quality at the target composition as well as at the edges of the allowable composition ranges and fulfillment of the end-of-shelf-life stability requirements of 36 months at the intended storage temperature (2°C-8°C). Extrapolation of 24 months' formulation robustness data to end of shelf life showed that the MAb formulation was robust within the claimed formulation composition ranges. Based on this case study, we propose that a formulation can be claimed as "robust" if all drug substance and drug product critical quality attributes remain within their respective end-of-shelf-life critical quality attribute–acceptance criteria throughout the entire claimed formulation composition range.

Introduction

The goal of pharmaceutical development, including formulation and manufacturing process development, is to consistently deliver the intended quality of the product within allowable ranges. Regulatory agencies and industry experts have suggested rigorous approaches to strengthen pharmaceutical development, manufacturing, and quality assurance.[1-3] In August 2002, the Food and Drug Administration launched a new initiative entitled "Pharmaceutical cGMPs for the 21st century: A risk based approach".[4] The aim of this initiative was to introduce innovation into pharmaceutical development. In 2004, the Food and Drug Administration released the Process Analytical Technology Guidance for Industry, which outlined in detail the future expectations of health authorities with regard to developing and implementing effective and efficient innovative approaches in pharmaceutical development.[5] The guidance advocates "building quality into products" by science and risk-based approaches and recognizes statistical experimental design as one of the tools that enable a scientific risk-based approach.

In 2008 and 2009, the International Conference on Harmonization (ICH) guidelines Q8 and Q8(R2) defined quality-by-design (QbD) as "a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management".[6] These guidelines state that "Critical formulation attributes and process parameters are generally

Abbreviations: AC, acceptance criterion; Buff, buffer concentration; CQA, critical quality attribute; CI, confidence interval; DoE, design of experiments; DP, drug product; EoS, end of shelf life; HMWS, high–molecular weight species; ICH, International Conference of Harmonization; LMWS, low–molecular weight species; MLR, multiple linear regression; Q^2, coefficient of prediction; QbD, quality-by-design; R^2, coefficient of determination; TR, target range.

Conflicts of interest: Medical writing and editorial assistance were provided by Anna Murphy of Fishawack Communications, funded by F. Hoffmann-La Roche Ltd. Christine Wurth, Hanns-Christian Mahler (at the time when the study was performed), and Michael Adler are employees of F. Hoffmann-La Roche Ltd. Barthelemy Demeule is an employee of Genentech Inc.

Current address for Hanns-Christian Mahler: Lonza AG, Drug Product Services, Basel, Switzerland.

This article contains supplementary material available from the authors by request or via the Internet at http://dx.doi.org/10.1016/j.xphs.2016.02.013.

* *Correspondence to:* Christine Wurth (Telephone: +41-61-6889135; Fax: +41-61-6883430).

E-mail address: christine.wurth@roche.com (C. Wurth).

http://dx.doi.org/10.1016/j.xphs.2016.02.013
0022-3549/

identified through an assessment of the extent to which their variation can have impact on the quality of the drug product" and define a design space as "the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality". Although design space has primarily been used in the context of pharmaceutical manufacturing processes, it can also be applied to formulation, as confirmed by a question and answers document that was published by the ICH Quality Implementation Working Group in 2011 (Q8/Q9/Q10 Q&A [R4]).[7] The document specifies that "it may be possible to develop a formulation (not component but rather composition) design space consisting of the ranges of excipient amounts and its physicochemical properties (…)" and states that, "the applicant should justify the rationale for establishing the (formulation) design space with respect to quality attributes such as bioequivalence, stability, manufacturing robustness etc." Although intentional movement within a formulation composition design space is not typically desired for a drug product (DP), the goal of pharmaceutical development is to deliver a robust formulation encompassing that product quality is warranted not only at target/label claims for composition but also within permitted ranges of the label composition (accounting, e.g., for manufacturing process-related variations).

This allowed formulation composition range is characterized during development by performing "formulation robustness" studies. The aim of these studies is to select a commercial formulation that is sufficiently robust within the acceptable ranges around the nominal label claim and to meet the shelf life stability requirements, which are typically 24-36 months for pharmaceutical products and at least 18 months at refrigerated conditions (2°C-8°C) for biopharmaceutical DPs.

QbD tools, especially design of experiments (DoEs), have been applied by the pharmaceutical industry to assess and ensure formulation robustness.[8] Examples of formulation robustness studies describing, for example, the effect of excipient composition on process performance can be found in the literature. [9-13] The outcome of such robustness DoE studies can be judged on 2 criteria: (1) Is the resulting regression model statistically significant? (2) Are the output parameters (quality attributes) inside or outside predefined limits? The first criterion, "statistical significance," is routinely assessed when analyzing formulation robustness DoEs by statistical analyses such as multiple linear regression (MLR) methods. To fulfill the second criterion, "acceptable limits" are needed when assessing "robustness."[9-13]

Although this recent literature describes several case studies of formulation robustness, 2 challenges remain: first, the definition and setting of "acceptable limits" are needed to assess formulation robustness with regard to stability until end of shelf life (EoS) within a QbD approach; and second, at the time of market application submission, there are often no EoS stability data available from the robustness study at the intended storage condition. Thus, conclusions with regard to "formulation robustness" are typically based on accelerated temperature and short-term stability testing and are, therefore, more qualitative than quantitative in nature.

For critical quality attributes (CQAs) related to formation of degradant(s) over shelf life, assessment of "acceptable limits" is challenging. Setting acceptable limits for CQAs in a QbD approach for biopharmaceuticals is described in more detail in a series of QbD articles that have been recently submitted for publication.[14-16] In brief, once the CQAs have been identified, an acceptable level or range for each CQA needs to be defined which is called the CQA—acceptance criterion (CQA-AC). The CQA-AC defines the product quality requirement for end of DP shelf life and is set for each CQA based on its potential impact on efficacy, immunogenicity, pharmacokinetics, and patient safety, considering, for example, relevant clinical experience, results from process characterization and validation studies, and clinical data from similar products, including data from literature. The CQA-AC is the sum of all allowed changes for 1 CQA with regard to drug substance (DS) and DP process and stability and can be further divided into allowable ranges for these 4 processing and storage steps, namely for the (1) DS manufacturing process, (2) DS storage ("DS allowable stability range"), (3) DP manufacturing process, and (4) DP storage ("DP allowable stability range"). The limits of each of these 4 process and storage steps are termed CQA target ranges (CQA-TRs). Generally, acceptable limits for process steps are assessed primarily based on characterization and validation data from at-scale data and scale-down models and are further narrowed by a 5% reduction of the derived limit to account for modeling uncertainties. The CQA-TRs for the DS and DP process usually correspond to the DS and DP CQA release specifications, respectively. The DP allowable stability range is calculated as the difference between the DP release specification and the CQA-AC at EoS.

In this case study, the following DP allowable stability ranges were defined based on assessments including product-specific relevant clinical experience, results from process characterization and validation studies, and clinical data from similar products (including data from literature) that are beyond the scope of this article: 13.1 area % for acidic variants, 0.8 area % for high—molecular weight species (HMWS), 1.6 area % for low—molecular weight species (LMWS), and 7.8 area % for oxidized variants. The described DP allowable stability ranges for MAb-A CQAs are very product-specific and need to be determined case by case for each product based on the previously mentioned process.

In principle, formulation robustness can be determined by analyzing the respective quality attributes in real time until EoS is reached. In practice, scientists involved in formulation and DP development wish to obtain an indication of formulation robustness without waiting an unrealistic amount of time. This is comparable to claiming a shelf life for a formulation based on statistical analysis of short-term stability data.[17,18] We propose that this concept may be applied to an entire formulation composition range. Our case study addresses the challenge of how to quantitatively assess formulation robustness for a biotechnology product before real-time EoS data are available by combining concepts of ICH Q8 Pharmaceutical Development ("formulation composition design space") and Q1A(R2) Stability Testing of New Drug Substances and Products ("predicting a shelf life").[6,7,18] Our primary interest with regard to formulation robustness is to guarantee stability of the formulation during storage and until EoS.

Where predefined acceptance limits are available, in general, the following 4 formulation robustness study outcomes can be distinguished: (1) nonsignificant statistical model and all output parameter values (CQA values in our case) are clearly inside predefined limits; (2) significant statistical model and all output parameter values are clearly inside predefined limits; (3) significant statistical model and output parameter values are outside predefined limits; and (4) nonsignificant statistical model and output parameter values are outside predefined limits. The first case provides little knowledge about how formulation factors affect product performance. However, this case represents the ideal outcome of a formulation robustness study because it reflects that the CQA response is not significantly affected by the input (formulation) parameter variations, thus confirming a robust formulation within the predefined limits. The second case also confirms formulation robustness within predefined acceptance criteria. In addition, the statistically significant model would show how formulation factors affect product performance and allows extrapolation according to Q1A if end-of-shelf-life data are not available at time of analysis. The third case provides information on how formulation factors

affect product performance but shows that the formulation is not robust within the tested formulation composition range. As the regression model is statistically significant, it allows the identification of formulation parameters or formulation parameter ranges that are accountable for the values outside the specified limits. Furthermore, the model can be used to predict an acceptable formulation parameter range where all the values for the assessed quality attributes will be inside the specified limits, thus supporting formulation robustness by adjustment of, for example, formulation parameter ranges. As for the second case, if EoS data are not available at time of analysis, extrapolation, for example, according to ICH Q1A, is necessary. Finally, the fourth case provides little formulation knowledge, and formulation robustness cannot be claimed. No statistically significant regression model exists, therefore the acceptable formulation parameter ranges cannot be set based on the existing data, and additional experiments are required to identify the influencing formulation parameters.

Here, we illustrate cases 1-3 with examples from a formulation robustness study performed for a model MAb-A before biologic license application/marketing authorization application submission. Additionally, a quality attribute showing nonlinear degradation kinetics ("HMWS") will be described.

Materials and Methods

Materials

An IgG1 MAb provided by F. Hoffmann-La Roche Ltd. (Basel, Switzerland) was used for these studies. Histidine free base and histidine monohydrochloride were purchased from Ajinomoto (Raleigh, NC). Trehalose was purchase from Ferro Pfanstiehl (Waukegan, IL) and Poloxamer 188 from BASF (Florham Park, NJ). Dulbecco's phosphate-buffered saline 10× was obtained from GIBCO (Invitrogen, San Diego, CA). All other chemicals used were of analytical grade and obtained from commercial sources.

DoE Robustness Study Design

A multivariate stability study was performed with 3 formulation factors at 2 levels, namely MAb-A concentration (MAb), solution pH (pH), and buffer concentration (Buff). Two additional factors, namely tonicifier and surfactant concentration, were tested in a previous short-term robustness DoE study (generation of up to 6 months stability data) that included varying pH levels, surfactant, tonicifier, and MAb-A concentrations (data not shown). As the 2 formulation factors surfactant and tonicifier were found to have no significant impact on MAb-A CQAs and no interaction with other factors in the ranges that were tested, these 2 parameters were excluded from the present study design and kept at constant concentration levels in all tested formulations.

The 3 factors MAb-A concentration, pH, and buffer concentration were investigated in a full factorial design resulting in 8 experiments and 3 center points, the latter corresponding to the target formulation, as shown in Table 1. The applied design allows estimation of all linear terms and all interaction terms. The tested formulation parameter ranges were defined to cover either specification acceptance criteria or weighing tolerances of manufacturing when adding formulation excipients (e.g., compounding). The experimental design plan is provided in Table 1, and the actual design is shown in Figure 1. The quality of MAb-A in the formulation compositions described in Table 1 was evaluated after 0, 13, 26, 40, 52, and 104 weeks of storage at 2°C-8°C by size exclusion high-performance liquid chromatography, ion exchange high-performance liquid chromatography, and protein A chromatography as described in later section. In addition, formulations were analyzed after storage at 25°C for up to 40 weeks and after storage at 40°C for up to 13 weeks (data not shown).

Table 1
Compositions of MAb-A Formulations[a] Evaluated in a Full Factorial Design of Experiments

Sample Code	MAb-A Concentration (mg/mL)	pH	Buffer Concentration (mM)
F01	22.5	5.5	15
F02	27.5	5.5	15
F03	22.5	6.5	15
F04	27.5	6.5	15
F05	22.5	5.5	25
F06	27.5	5.5	25
F07	22.5	6.5	25
F08	27.5	6.5	25
F09 (Target)	25.0	6.0	20
F10 (Target)	25.0	6.0	20
F11 (Target)	25.0	6.0	20

[a] All formulations additionally containing 0.02% (wt/vol) surfactant and 240-mM tonicifier.

Manufacturing of Formulations

MAb-A formulations according to Table 1 were prepared by addition of trehalose and poloxamer 188 stock solution to histidine-buffered MAb-A solutions (pH: 5.5, 6.0, and 6.5) as required. Final formulations were filtered using 0.2-mm Sterivex GV (polyvinylidene fluoride) filters (Millipore, Bedford, MA), filled into appropriately washed and sterilized 6 mL Ø 20 mm glass type 1 vials, sealed with rubber stoppers, and crimped with aluminum caps. The 3−center point target formulations F9-F11 were manufactured as 3 individual formulation batches. Formulations were then stored for up to 104 weeks at a temperature of 2°C-8°C. In addition, formulations were stored at 25°C for up to 40 weeks and at 40°C for up to 13 weeks (data not shown).

Multivariate Data Analysis

Data analysis was performed using ORIGIN 7.5 (OriginLab) and MODDE, version 10.0 (Umetrics), software. MLR was used to estimate the effect of protein concentration (MAb), solution pH (pH), and buffer concentration (Buff) on the MAb-A quality attributes HMWS as detected by size exclusion high-performance liquid chromatography, acidic variants as detected by ion exchange high-performance liquid chromatography, and oxidized variants as detected by protein A chromatography. Visible and subvisible particles were also monitored in all samples, but as no significant changes over time were observed, no statistical data analysis was performed for these CQAs (data not shown).

Experimentally determined protein and histidine concentrations and pH levels were used for multivariate data analysis.

To focus solely on the impact of formulation parameters on DP stability before MLR analysis, that is, to focus on change over time, initial CQA values were subtracted from respective CQA values at later time points. MLR data sets for HMWS, acidic variants, and oxidized variants were analyzed separately with "time" included as a 4th input parameter in all regression models to allow for extrapolation until EoS. Nonsignificant terms were removed from the model for model optimization.

All experiments for a specific CQA were used to calculate a regression model, that is, in total, 88 data points per CQA regression model (11 formulations analyzed at 8 time points; initial analysis and after storage for 4, 8, 13, 26, 40, 52, 78, and 104 weeks, resulting in 88 data points per CQA). As we use MLR and least square analysis for data analysis, minimizing the sum of squared deviations

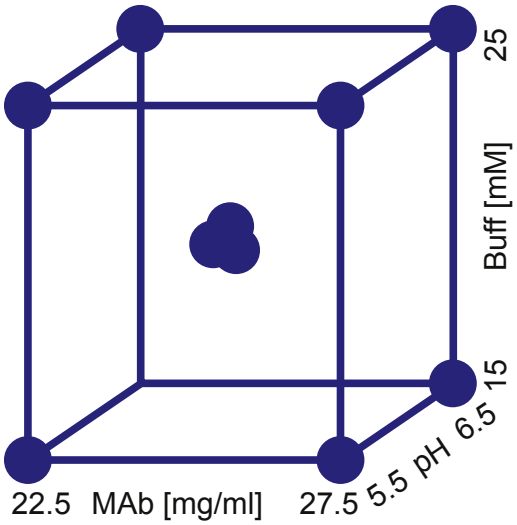

Figure 1. Design region of experimental plan. The multivariate stability study included 3 formulation factors at 2 levels, MAb-A concentration ("MAb"), solution pH ("pH"), and buffer concentration ("Buff"). These factors were investigated in a full factorial design resulting in 8 experiments and 3 center points.

between the observed and fitted values of each response, this has the advantage of the production of a number of useful model diagnostic tools such as the coefficient of determination (R^2) indicating the goodness of fit, that is, the proportion of variability explained by the model, and the coefficient of prediction (Q^2), a measure of how well the model will predict the responses for new experimental conditions. A model that explains the data well will have an R^2 and Q^2 close to 1.0. Furthermore, errors ("SDs") including the propagation of error are presented in the error bars (CIs) of coefficients and predictions. The model diagnostics for the tested product quality parameters are given in Supplementary Table 1 (Supporting Information).

Regression coefficient plots, displaying the respective regression coefficients of the fitted models with CIs, were used for model interpretation to evaluate the influence of the different formulation parameters on quality attributes. A regression coefficient is regarded as statistically significant if its CI is smaller than the coefficient. Ninety-five percent CIs were chosen for model predictions after Q1A(R2) Stability Testing of New Drug Substances and Products stating that: "An approach for analyzing data of a quantitative attribute that is expected to change with time is to determine the time at which the 95 one-sided confidence limit for the mean curve intersects the acceptance criterion."[18]

Analytical Methods

Analytical samples were analyzed once per analytical technique using validated and qualified methods and equipment (i.e., understanding the full method performance including precision, accuracy, and reproducibility) as described in the following text, following internal as well as ICH Q1A requirements for DP. True replicates were performed in the center points (target formulations F9, F10, and F11) which are added to the design as an estimation of experimental noise and to check for inherent variability of the system.

Size Exclusion High-Performance Liquid Chromatography

Size exclusion chromatography was performed on an Alliance 2695 HPLC system equipped with a 2487 UV detector (Waters Corporation, Milford, MA) and a Tosoh Bioscience TSK-Gel G3000SWXL column with an oven temperature set to 25°C. The mobile phase consisted of a 0.2-M dipotassium hydrogen phosphate/potassium dihydrogen phosphate buffer and 0.25-M potassium chloride (pH 7.0) at a flow rate of 0.5 mL/min for 30 min. Samples were diluted to 10 mg/mL protein concentration, and 150 μL was injected, corresponding to a loading amount of 150 μg. Detection was performed at a wavelength of 280 nm. Data analysis was performed using the software Empower 2 Chromatography Data System (Waters Corporation), and results were reported as area % HMWS and area % LMWS. The term HMWS describes all MAb-A size variants which elute earlier than the main peak, comprising dimers and larger soluble aggregates. The term LMWS describes all MAb-A size variants which elute later than the main peak.

Ion Exchange High-Performance Liquid Chromatography

Ion exchange chromatography was performed on an Alliance 2695 HPLC system equipped with a 2487 UV detector (Waters Corporation) and a Dionex ProPac WCX-10, 4 × 250 mm column with an oven temperature set to 40°C. The mobile phases consisted of 10-mM sodium phosphate (pH, 6.1; mobile phase A) and 10-mM sodium phosphate, 750-mM sodium chloride (pH, 6.1; mobile phase B). Samples were diluted to 0.5-mg/mL protein concentration, and 100 μL was injected, corresponding to a loading amount of 50 μg. The mobile phases were pumped at a flow rate of 1.0 mL/min in a gradient mode [(time in min vs. percentage of mobile phase B) = 0:5, 5:5, 35:20, 45:95, 47:95, 50:5, 60:5]. Detection was performed at a wavelength of 280 nm. Data analysis was performed using Empower 2 Chromatography Data System software (Waters Corporation), and results were reported as area % acidic variants. The term "acidic variants" describes all MAb-A charge variants which elute earlier than the main peak.

Analytical Protein A Chromatography

The presence of oxidized methionine residues in the antibody was quantified using analytical protein A chromatography as described previously. [19] In brief, the content of oxidized MAb-A variants was determined using an Alliance 2695 HPLC system equipped with an 2487 UV detector (Waters Corporation) and a Poros A/20 4.6 mm D × 50 mm L column. The mobile phases consisted of Dulbecco's phosphate-buffered saline 1× in water (pH, 7.4; mobile phase A) and 100-mM acetic acid with 150-mM sodium chloride in water (pH, 2.8; mobile phase B). Samples were diluted to 10-mg/mL protein concentration, and the injection volume was set to 100 μL. The mobile phases were pumped at a flow rate of 2.0 mL/min in a gradient mode [(time in min vs. percentage of mobile phase B) = 0:0, 40:60, 41:100, 51:100, 52:0, 62:0]. Detection was performed at a wavelength of 280 nm. Data analysis was performed using Empower 2 Chromatography Data System software (Waters Corporation), and results were reported as area % oxidized variants.

Protein Concentration

Protein concentration in solution was determined by photometric analysis of diluted samples at 280 nm according to Lambert-Beer law.

Histidine Concentration

The content of histidine was determined by reversed-phase HPLC using an Alliance 2695 HPLC system equipped with an 2487 UV detector (Waters Corporation) and an ODS-A 250 × 4.6 mm column with an oven temperature set to 25°C.

The mobile phase consisted of a 10-mM sodium phosphate buffer solution (pH 7.0) at a flow rate of 0.5 mL/min for 15 min. Samples were diluted 1:10 with mobile phase, and 10 μL was injected. Detection was performed at a wavelength of 210 nm. Data

analysis was performed using Empower 2 Chromatography Data System software (Waters Corporation).

Solution pH

Formulation solution pH values were determined using a Metrohm 781 pH meter and a Biotrode micro pH glass electrode (Metrohm AG, Herisau, Switzerland).

Results

Case 1: LMWS

It can be assumed from the initial assessment of LMWS data that this response is robust because all measured values are clearly inside the specifications, that is, below 1.6 area %, and no major variation in LMWS over time is seen (Fig. 2).

Case 2: Acidic Variants

The response "acidic variants" is an example where a statistically significant regression model exists, and all responses are well inside predefined specification limits when extrapolating time to EoS.

The raw data scatter plot for the data set "acidic variants" including specification limits is shown in Figure 3. The obtained regression model was significant with an R^2 of 0.86 and a Q^2 of 0.84. The coefficient plot shows 2 significant linear coefficients, pH and time (Fig. 3). The positive orientation of the coefficient bar for pH indicates that a high pH will lead to a higher increase in acidic variants during storage at intended storage conditions (2°C-8°C) compared to formulations with a solution pH at the lower edge of the formulation composition range. The interaction term pH × time originates from the study setup where all output parameters are normalized to 0 at start of the study to focus on stability impact assessment. Any effect of pH will then increase over time giving the interaction term.

Extrapolation up to 3 years for the worst case formulation including the 95% CI resulted in a predicted value of 5.7 area %, which is clearly below the predefined limit of 13.1 area %. It can be assumed from this assessment that this response is robust (Table 2).

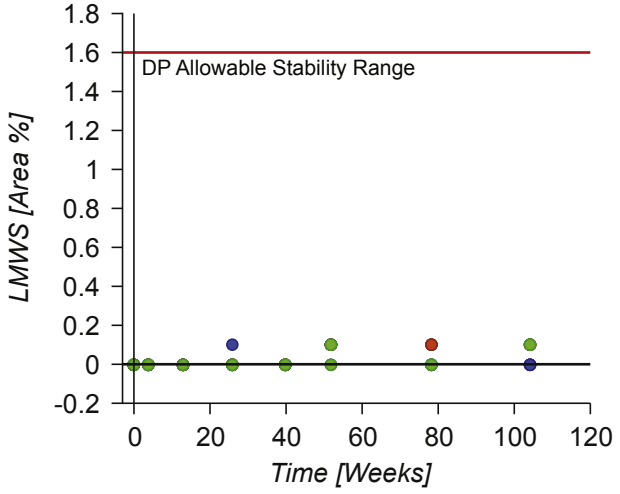

Figure 2. Raw data scatter plot for low–molecular weight species. Data points colored by formulation pH: red circles pH 6.5, green circles pH 6.0, blue circles pH 5.5. It shows all formulations at each time point; however, the data points for different formulations overlap because all values for LMWS are either 0 or 0.1 area %.

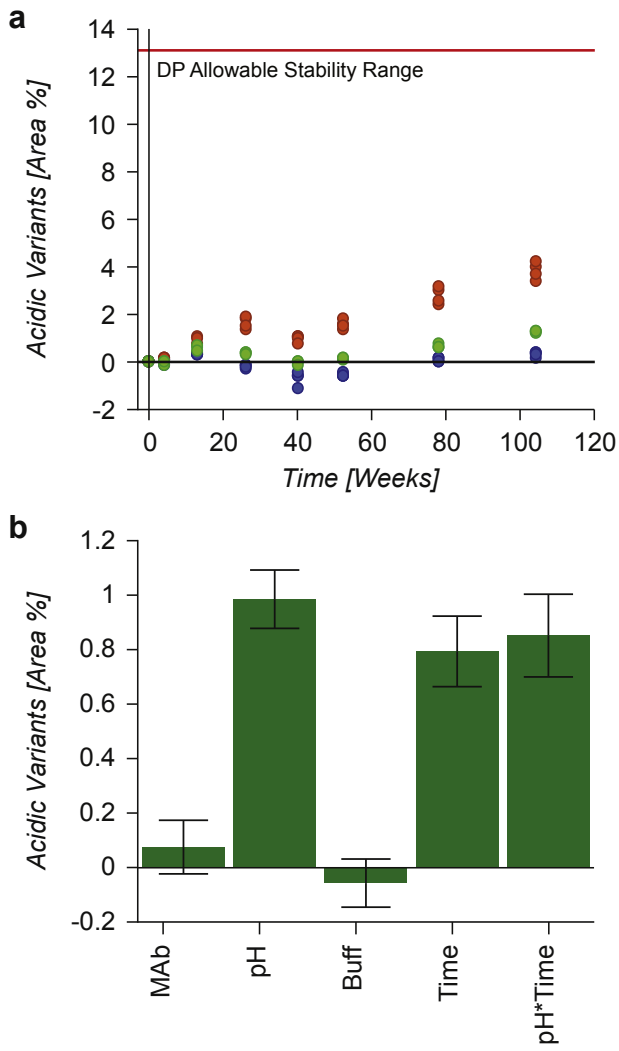

Figure 3. Acidic variants. (a) Raw data scatter plot for acidic variants. Data points colored by formulation pH: red circles pH 6.5, green circles pH 6.0, blue circles pH 5.5. (b) Coefficient plot for acidic variants showing 2 significant linear coefficients, pH and time. The positive orientation of the coefficient bar for pH indicates that a high pH will lead to a higher increase in acidic variants during storage at 2°C-8°C.

Case 2: HMWS

As for acidic variants, the response "HMWS" is an example where a statistically significant regression model exists, and all response are well inside predefined specification limits. However, as shown in Figure 4, MAb-A aggregation at 2°C-8°C follows a nonlinear kinetic; linearization of the data is thus needed before performing data analysis by MLR and extrapolating to EoS.

It is known that aggregation kinetics are highly complex, and it has been shown that logistic functions empirically fit a wide variety of protein aggregation curves. [20,21] In our case, we assumed a simplified model, a pseudo first order kinetic reaction for HMWS increase over time (Eq. 1), with y = HMWS [area %], k = first order reaction constant, and t = time [weeks], and we obtained a good fit to the experimental data as well as a good regression model.

$$y = a\left(1 - e^{-kt}\right) \quad (1)$$

Linearization of the HMWS data set can then be achieved by plotting "Etime" = 1−EXP(−k_n × time) (instead of "time") versus

Table 2
Predicted Values for Acidic Variants at the 156-Week (36-Month) Time Point With 95% Upper and Lower CIs

Design of Experiments Study Setup	Input Parameters				Output Parameter (Area %)		
	Mab-A (mg/mL)	pH	Buffer (mM)	Time (wk)	Acidic Variants	95% Lower	95% Upper
F09 (Target formulation)	25.0	6.0	20	156	2.4	2.1	2.7
F01	22.5	5.5	15	156	−0.3	−0.8	0.2
F02	27.5	5.5	15	156	−0.2	−0.6	0.3
F03	22.5	6.5	15	156	5.1	4.6	5.6
F04	27.5	6.5	15	156	5.2	4.7	5.7
F05	22.5	5.5	25	156	−0.4	−0.9	0.1
F06	27.5	5.5	25	156	−0.3	−0.8	0.2
F07	22.5	6.5	25	156	5.0	4.5	5.5
F08	27.5	6.5	25	156	5.1	4.6	5.6

All predicted upper values at the 156-week time point are well within the allowable drug product stability target range of 13.1 area % for acidic variants. Negative values are within the validated method variation of the applied analytical method (ion exchange high-performance liquid chromatography [IE-HPLC]).

"HMWS," with k_n = first order reaction constants for formulations n = 1-11.

The k_n values for the 11 robustness formulations were obtained by nonlinear curve fitting using the Origin software and fitting the individual formulation HMWS data sets to the equation $y = a_n(1-EXP(-k_n \times t)$, describing first order kinetics for increase of a product. 'Etime' = $1-EXP(-k_n \times$ time) was then used instead of time as a fourth input parameter for the regression model.

All statistically significant linear and interaction coefficients, also including small ones, were included into the regression model used for predictions. The resulting MLR model was statistically significant with an R^2 of 0.97 and a Q^2 of 0.96 and revealed time as the main driver for changes in the CQA value (Fig. 4) with a respective coefficient being >7 times larger than the other coefficients. The 3 other linear coefficients, namely MAb, pH, and Buff were also statistically significant (i.e., a higher pH or protein concentration leads to a higher increase in HMWS, whereas a higher buffer concentration leads to a lower increase in HMWS during storage), but all were small and close to statistically insignificant so their contribution to variations in the output parameters during storage at 2°C-8°C is of no practical relevance. However, although out of scope for this article, their effects should not be ignored because they might become significant in the case of temperature excursions. Several of the interactions also appear to be statistically significant and were included into the regression model. One might speculate about the nature of these coefficients; however, interaction coefficients are by nature not easily explainable in complex systems, and furthermore, the model for increase in HMWS over time is not linear, complicating even more the interpretation of these interactions.

All responses up to the 24-month time point were well below the defined acceptance criteria limit of 0.8 area % and thus within specifications. As the regression model for HMWS was statistically significant, model prediction tools can be used to extrapolate the input variable "Etime" to EoS (156 weeks). Table 3 depicts predicted values for HMWS at the 156-week (36-month) time point with 95% CI for the tested formulation composition range. Result shows that predicted values for HMWS at EoS range from 0.26 to 0.43 area % considering the upper 95% CI (Table 3), which is still well below the specified limit of 0.8 area % at EoS for this CQA.

The formulation can be thus considered to be robust with regard to the CQA HMWS within the tested formulation composition range.

Case 3: Oxidized Variants

The response "oxidized variants" is an example where a statistically significant regression model exists, but this response will be outside predefined specification limits when extrapolating time to EoS.

The raw data scatter plot for the data set "oxidized variants" including predefined acceptance criteria limits (7.8 area % DP allowable stability range) is shown in Figure 5. The calculated

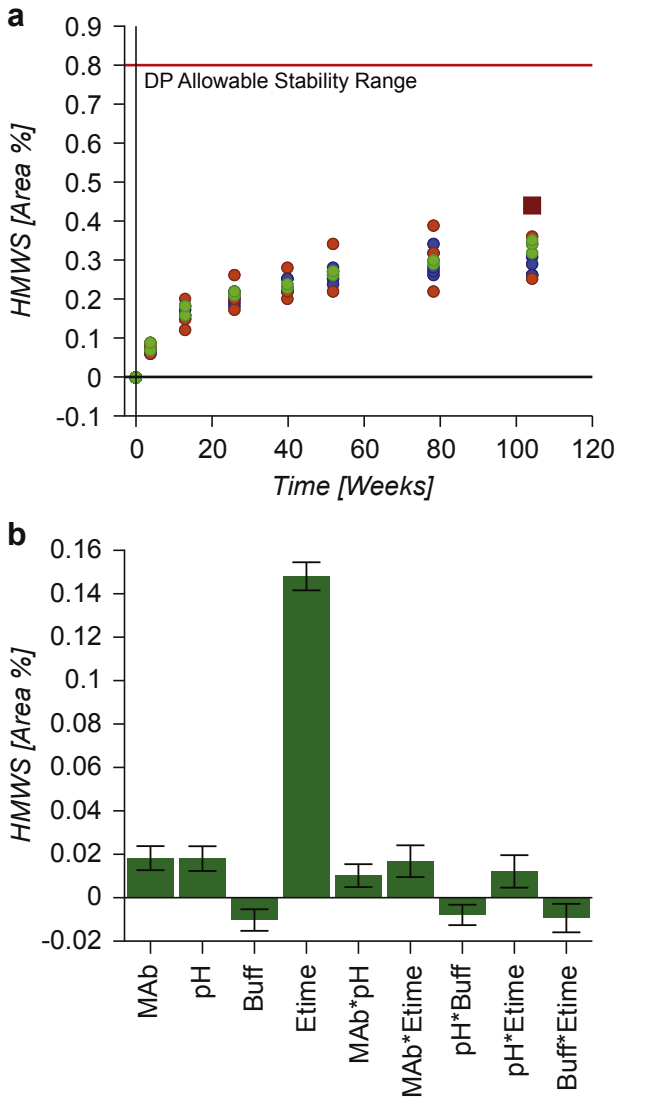

Figure 4. High–molecular weight species. (a) Raw data scatter plot for high–molecular weight species. Data points colored by formulation pH: red circles pH 6.5, green circles pH 6.0, blue circles pH 5.5. The red square represents formulation no. 4 which does not seem to follow the overall model. (b) Coefficient plot for high–molecular weight species showing time as the main driver for changes in the CQA value.

Table 3
Predicted Values for HMWS at the 156-Week (36-Month) Time Point With 95% Upper and Lower CIs

Design of Experiments Study Setup	Input Parameters				Output Parameter (Area %)		
	Mab-A (mg/mL)	pH	Buffer (mM)	Time (wk)	HMWS	95% Lower	95% Upper
F09 (Target formulation)	25.0	6.0	20	156	0.31	0.30	0.32
F01	22.5	5.5	15	156	0.27	0.25	0.28
F02	27.5	5.5	15	156	0.32	0.30	0.33
F03	22.5	6.5	15	156	0.32	0.30	0.34
F04	27.5	6.5	15	156	0.41	0.39	0.43
F05	22.5	5.5	25	156	0.24	0.23	0.26
F06	27.5	5.5	25	156	0.29	0.28	0.31
F07	22.5	6.5	25	156	0.27	0.25	0.28
F08	27.5	6.5	25	156	0.35	0.34	0.37

All predicted upper values at the 156 week time point are well within the allowable drug product stability target range of 0.8 area % for HMWS.

regression model was significant with an R^2 of 0.92 and a Q^2 of 0.91. The coefficient plot (Fig. 5) depicts time and pH as statistically significant linear coefficients. The positive orientation of the coefficient bar for pH indicates that a high pH will lead to a higher increase in oxidized variants during storage at 2°C-8°C compared to formulations with a solution pH at the lower edge of the formulation composition range.

All responses up to the 24-month time point were still below the allowable DP stability range of 7.8 area % oxidized variants. However, when using model prediction tools to extrapolate the input variable "time" to EoS (36 months) for worst case high-pH formulations, the predicted CQA values at EoS exceeded the allowed stability TR (Table 4). This indicates that the formulation is not sufficiently robust at the edges of the proposed formulation composition range.

As the regression model is statistically significant, it can be used to adjust the composition range by predicting an acceptable input (formulation) parameter range, where all the values for the output parameter will be inside the specification limit. In this case, narrowing the formulation composition range, that is, narrowing the upper pH specifications from 6.5 to 6.3 will reduce the predicted 95% upper value for oxidized variants at EoS to 7.3 area % (Table 4) and thus will then result in a robust formulation.

Discussion and Conclusions

Despite the strong recommendation from regulatory agencies that risk analysis as well as DoEs should be used to better understand formulations and processes, there are no published examples of a complete case study outlining the entire QbD procedure for defining a robust formulation from concept to conclusion.

Gabrielsson et al.[9] (2008) evaluated the robustness of 2 tablet formulations and concluded that a system is robust if the CQAs vary within acceptable limits as critical process parameters are changed over normal variation intervals. Later, Awotwe-Otoo et al.[10] (2012) used QbD approaches to evaluate the effect of several formulation variables and their interactions on quality attributes of a lyophilized monoclonal antibody. They also identified optimal concentrations of excipients and pH to define a design space.

Although many of the published case studies on formulation robustness are linked to process performance, there are few publications about formulation robustness in the context of product stability throughout the intended DP shelf life, especially for biotech products. An attempt to discuss formulation design and effects of variations in the formulation composition on stability was carried out in 2009 in the A-Mab case study. [22] A set of DoE studies was used to verify a formulation composition range. It was observed that aggregation increased above pH 5.3 for storage at 40°C and was dependent on protein concentration. The author's conclusion was qualitative rather than quantitative with regard to formulation robustness over shelf life: "Due to the dependence of aggregation on pH, the formulation range for pH should not exceed 5.6".

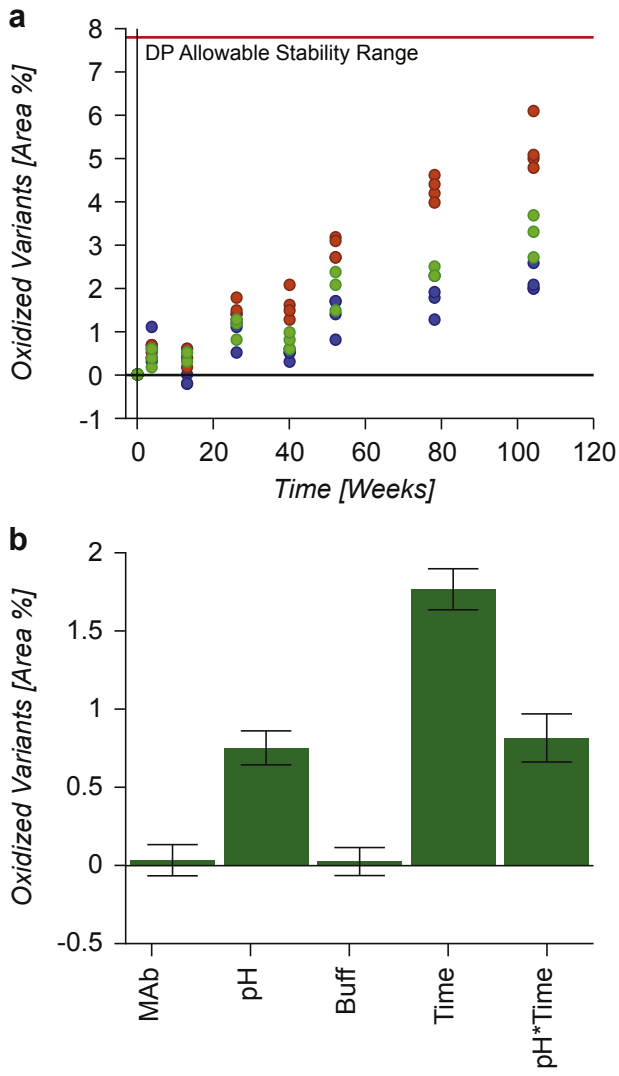

Figure 5. Oxidized variants. (a) Raw data scatter plot for oxidized variants. Data points colored by formulation pH: red circles pH 6.5, green circles pH 6.0, blue circles pH 5.5. (b) Coefficient plot for oxidized variants depicting pH and time as statistically significant linear coefficients. The positive orientation of the coefficient bar for pH indicates that a high pH will lead to a higher increase in oxidized variants during storage at 2°C-8°C.

Table 4
Predicted Values for Oxidized Variants at the 156-Week (36-Month) Time Point With 95% Upper and Lower CIs

Design of Experiments Study Setup	Input Parameters				Output Parameter (Area %)		
	MAb-A (mg/mL)	pH	Buffer (mM)	Time (wk)	Oxidized Variants	95% Lower	95% Upper
Original formulation composition range							
F09 (Target)	25	6.0	20	156	5.4	5.1	5.7
F01	22.5	5.5	15	156	3.0	2.5	3.5
F02	27.5	5.5	15	156	3.0	2.5	3.5
F03: Exceeds limit	22.5	6.5	15	156	7.8	7.3	8.2
F04: Exceeds limit	27.5	6.5	15	156	7.8	7.3	8.3
F05	22.5	5.5	25	156	3.0	2.5	3.5
F06	27.5	5.5	25	156	3.1	2.6	3.6
F07: Exceeds limit	22.5	6.5	25	156	7.8	7.3	8.3
F08: Exceeds limit	27.5	6.5	25	156	7.9	7.4	8.4
Adjusted range							
Narrowed pH range	27.5	**6.3**	25	156	6.9	6.5	7.3
Shortened shelf life	27.5	6.5	25	**145**	7.3	6.9	7.8

The predicted upper values for worst case formulations F03/F04 and F07/F08 (high pH) exceed the allowable drug product stability target range of 7.8 area %. Adjusting the claimed formulation composition range, that is, narrowing the pH range from 6.5 to 6.3 or shortening shelf life to 145 weeks (values in bold) will result in predicted upper values for oxidized variants within the specified limits.

In 2011, an article by Martin-Moe et al.[11] described a formulation robustness case study where a DoE was set up to investigate formulation robustness by testing the effects of pH, concentration of surfactant, buffer, tonicifying agent, as well as time and temperature. Stability data from a 6-month isothermal study showed that pH, time, and temperature had a notable impact on CQAs at 25°C. The article concluded that based on these results, combined with additional univariate and multivariate studies, "CPPs are identified (pH, time, and temperature), multivariate acceptable ranges are defined, and a manufacturing target is recommended for the formulation".

A comprehensive article regarding formulation design space and robustness was published by Grillo et al. in 2010.[13] The article presents helpful considerations for performing robustness studies as well as 2 case studies in which DoE was used to determine the robustness of protein formulations to changes in protein, excipient, and pH levels. The case studies illustrate 2 outcomes that can be obtained from a formulation robustness study: first, when there is no effect of any formulation factor, it is straightforward to conclude that the formulation is robust to the tested variations in excipient and protein levels; and second, when there are effects, acceptance criteria need to be defined so that the quality of the DP is not affected beyond an acceptable level. For example, the results from the excipient robustness study in case 2 "prompted well-defined pH acceptance criteria in buffer preparation records and on the product specification", acknowledging the need of "acceptable limits" to quantitatively assess formulation robustness. However, no further guidance was given on how to define such "acceptable acceptance criteria" and "acceptable levels" for product quality throughout shelf life.

Recently, Sreedhara et al. published an overview on applying QbD principles to formulation development of biopharmaceutical products, including 3 case studies on formulation robustness.[12] The authors emphasize that robust protein formulation development is the keystone for all biopharmaceuticals and provide helpful examples of risk assessment tools to identify formulation components to be tested, DoE study setups, sampling plans, and data evaluation for a formulation robustness study.

In our study, first, we predefined acceptance limits for the CQAs tested in formulation robustness studies, based on QbD principles and in alignment with the project specific overall control strategy. Second, an MLR model was calculated. In the case that the MLR model is statistically not significant, we can deduce that the formulation is robust within the formulation composition ranges. In the case that the MLR model is statistically significant, we extrapolate to EoS and assess whether CQAs remain within those predefined limits; if we stay within the limits, we can infer that the formulation is robust within the formulation composition ranges, and if we exceed the limits, we can use the MLR model to adjust the formulation composition ranges accordingly, for example, by narrowing formulation composition ranges (e.g., narrowing pH specifications) or, possibly, by widening DP CQA allowable stability ranges or acceptance criteria, if warranted and justified.

Two important considerations, when using our proposed QbD procedure for defining a robust formulation, are that CQAs and CQA-ACs need to be defined early on, as a quantitative result cannot be generated without a DP stability range. Furthermore, there may be a need to change DP stability ranges as new information becomes available, which is usually the case during development, where control strategy and specifications can also evolve and if justified by scientifically valid reasons, these changes may be permissible.

The presented case study proposes a quantitative solution that demonstrates how to assure formulation robustness until EoS, based on real-time stability data at intended storage temperature and inclusion of time as a 4th input parameter in all regression models, if no EoS data are available at the time of biologic license application and marketing authorization application submission. Extrapolation to 36 months for predicting shelf life from both periods of 0-12 months and 0-24 months was evaluated, and both extrapolations predicted similar results with regard to degradation product content at EoS.

Although formulation parameter effects might be close to statistically insignificant to be of practical relevance at refrigerated storage temperatures, their effects should not be ignored because temperature excursions may exacerbate the effects of these parameters. In principle, the same concept described in this case study for real-time 2°C-8°C data can be used for evaluating temperature excursions at elevated temperatures, and all MAb-A formulations in the presented study design were indeed also stored at elevated temperature (25°C) for up to 40 weeks and at accelerated temperature (40°C) for up to 13 weeks (data not shown). It would also be interesting to evaluate whether the data set at accelerated storage conditions can achieve the same information as the data set for long-term storage at 2°C-8°C; however, relating the available data sets to the stability behavior of MAb-A at 2°C-8°C is outside the scope of this article but might be the topic of a further publication.

Formulation robustness studies have a dual purpose of supporting both range settings for formulation parameters as well as justification of specifications for CQAs. An advantage of developing an MLR model for a formulation robustness data set, including time as an input parameter, is that the MLR model can be used to make predictions within the data set to adjust the allowed formulation composition range or narrow shelf life specifications if necessary.

For a linear data set, the preferred methodology is MLR. For nonlinear data, such as HMWS, some type of linearization is necessary before an MLR model can be calculated. This should, however, be based on and be in line with mechanistic understanding of the underlying degradation pathways. However, if such understanding is not fully warranted, then statistical methods can be used to investigate how well the data can be fitted to a simplified model. A detailed scientific explanation about the nature of the interactions in our HMWS case would need more experiments which is out of scope for this article but is in consideration for a further study, focusing on the underlying mechanistics of MAb-A aggregation.

A potential problem can arise if one formulation does not follow the overall model. In this case, the predictions obtained are less precise, and consequently, a larger safety margin (CQA-TR) between the predicted value and the respective limit is required. In the presented case study, this problem potentially occurs for formulation no. 4 in the HMWS example (see red square in Fig. 4a). In such a case, a sufficiently large CQA TR needs to be chosen.

Based on this case study, we propose that a formulation can be claimed as robust if all DS and DP CQAs remain within their respective EoS CQA-AC throughout the entire claimed formulation composition range.

Acknowledgments

The authors thank Dr. Erik Johansson, Umetrics AB, for his strong support with MODDE multivariate data analysis and setting up multiple linear regression models for extrapolation of stability data within a formulation composition range, Eva Keller and Juliana Ott for their assistance with the robustness study laboratory work, and Yvonne Lentz, Amanda Pace, and Sreedhara Alavattam for fruitful discussions about formulation robustness.

This study was funded by F. Hoffmann-La Roche Ltd.

References

1. Rathore AS. Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends Biotechnol*. 2009;27(9):546-553.
2. Rathore AS, Winkle H. Quality by design for biopharmaceuticals. *Nat Biotechnol*. 2009;27(1):26-34.
3. Kozlowski S, Swann P. Current and future issues in the manufacturing and development of monoclonal antibodies. *Adv Drug Deliv Rev*. 2006;58(5-6):707-722.
4. US Food and Drug Administration. *Pharmaceutical cGMPs for the 21st Century: A Risk Based Approach*; 2004. Available at: http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/ Manufacturing/QuestionsandAnswersonCurrentGoodManufacturingPracticescGMPforDrugs/UCM176374.pdf. Accessed March 14, 2016.
5. US Food and Drug Administration. *PAT Guidance for Industry—A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance*; 2004. Available at: http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf. Accessed March 14, 2016.
6. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). *ICH Harmonized Tripartite Guideline: Q8(R2) Pharmaceutical Development*; 2009. Available at: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm073507.pdf. Accessed March 14, 2016.
7. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). *Guidance for Industry: Quality Implementation Working Group on Q8, Q9, and Q10 Questions and Answers(R4)*; 2011. Available at: http://www.fda.gov/downloads/Drugs/.../Guidances/UCM210822.pdf. Accessed March 14, 2016.
8. Montgomery D, ed. *Design and Analysis of Experiments*. 7th ed. Hoboken, NJ: John Wiley & Sons; 2008.
9. Gabrielsson J, Sjostrom M, Lindberg NO, Pihl AC, Lundstedt T. Robustness testing of a tablet formulation using multivariate design. *Drug Dev Ind Pharm*. 2006;32(3):297-307.
10. Awotwe-Otoo D, Agarabi C, Wu GK, et al. Quality by design: impact of formulation variables and their interactions on quality attributes of a lyophilized monoclonal antibody. *Int J Pharm*. 2012;438(1-2):167-175.
11. Martin-Moe S, Lim FJ, Wong RL, Sreedhara A, Sundaram J, Sane SU. A new roadmap for biopharmaceutical drug product development: integrating development, validation, and quality by design. *J Pharm Sci*. 2011;100(8):3031-3043.
12. Sreedhara A, Lentz Y, Schoenhammer K, Stark C. Application of QbD principles to late-stage formulation development for biological liquid products. In: Jameel F, Hershenson S, Khan MA, Martin-Moe S, eds. *Quality by Design for Biopharmaceutical Drug Product Development. AAPS Advances in the Pharmaceutical Sciences Series 18*. New York, NY: Springer; 2015:115-135.
13. Grillo AO, Kane M, Penn N, Perkins M. Characterizing the formulation design space. *Biopharm Int*. 2010;23(3):30-39.
14. Alt N, Zhang TY, Motchnik P, et al. Determination of critical quality attributes for monoclonal antibodies using quality by design principles. *Biologicals*. 2015 [in production].
15. Hakemeyer C, McNight N, St. John R, Meier S, Trexler-Schmidt M, Kelley B, Zettl F, Puskeiler R, Kleinjans A, Lim F, Wurth C. *Process Characterization and Design Space Definition*. 2015. Manuscript submitted for publication.
16. Kepert F, Cromwell M, Engler N, Finkler C, Gellermann G, Gennaro L, Harris R, Iverson R, Kelley B, Krummen L, McKnight N, Motchnik P, Schnaible V, Taticek R. *Establishing a Control System Using QbD Principles*. 2015. Manuscript submitted for publication.
17. Waterman K. Understanding and predicting pharmaceutical product shelf-life. In: Huynh-Ba K, ed. *Handbook of Stability Testing in Pharmaceutical Development*. New York: Springer Science + Business Media; 2009. ed.
18. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). *ICH Harmonized Tripartite Guideline: ICH-Q1A(R2) Stability Testing of New Drug Substances and Products*; 2003. Available at: http://www.fda.gov/downloads/drugs/ guidancecomplianceregulatoryinformation/guidances/ucm073369.pdf. Accessed March 14, 2016.
19. Loew C, Knoblich C, Fichtl J, et al. Analytical protein A chromatography as a quantitative tool for the screening of methionine oxidation in monoclonal antibodies. *J Pharm Sci*. 2012;101(11):4248-4257.
20. Morris AM, Watzky MA, Finke RG. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim Biophys Acta*. 2009;1794(3):375-397.
21. Roberts CJ. Non-native protein aggregation kinetics. *Biotechnol Bioeng*. 2007;98(5):927-938.
22. CMC Biotech Working Group 2009. A-MAb: A Case Study in Bioprocess Development. Available at: https://www.casss.org/associations/9165/files/A-Mab_Case_Study_Version_2-1.pdf. Accessed March 14, 2016.