Protein Crystal Shape and Size Control in Batch Crystallization: Comparing Model Predictive Control with Conventional Operating Policies

Joseph Sang-II Kwon,[†] Michael Nayhouse,[†] Panagiotis D. Christofides,[‡] and Gerassimos Orkoulas*,[†]

[†]Department of Chemical and Biomolecular Engineering and [‡]Department of Electrical Engineering, University of California, Los Angeles, California 90095-1592, United States

ABSTRACT: In this paper, we focus on a batch protein crystallization process used to produce tetragonal hen egg white lysozyme crystals and present a comparative study of the performance of a model predictive control (MPC) strategy formulated to account for crystal shape and size distribution with conventional operating strategies used in industry, namely, constant temperature control (CTC) and constant supersaturation control (CSC). Initially, a comprehensive, batch crystallizer model is presented involving a kinetic Monte Carlo (kMC) simulation model which describes the nucleation and crystal growth via adsorption, desorption, and migration mechanisms on the (110) and (101) faces and mass and energy balances for the continuous phase, which are developed to estimate the depletion in the protein solute concentration and the variation in the crystallizer temperature. Existing experimental data are used to calibrate the crystal growth rate and to develop an empirical expression for the nucleation rate. Simulation results demonstrate that the proposed MPC, adjusting the crystallizer jacket temperature, is able to drive the crystal shape to a desired set-point value with a low polydispersity for crystal size compared to CTC and CSC operating policies, respectively. The proposed MPC determines the optimal operating conditions needed to obtain protein crystals of a desired shape and size distribution as it helps avoid the small crystal fines at the end of the batch run.

INTRODUCTION

Protein crystallization plays a crucial role in the $1 trillion pharmaceutical industry and has been a major contributor to both scientific advancement and economic growth. More than 100 therapeutic proteins have been licensed and a number of additional therapeutic proteins are currently under testing. For example, therapeutic proteins such as albumin which regulates the colloidal osmotic pressure of blood and globulin which boosts the immune system in our body against infectious diseases are the main proteins of human cells. Additionally, the therapeutic protein structure can be investigated via nuclear magnetic resonance for proteins of small molar mass (less than 30 000). Alternatively, X-ray crystallography is suggested for the study of the structure of proteins with molecular weight over 30 000.[1−3] To use these methods, protein crystals, which are usually produced through a batch crystallization process, need to be of desired morphology with a low polydispersity. Here, we use tetragonal hen egg white (HEW) lysozyme which is a widely used model protein and relatively easily crystallizable with a molecular weight of 14 388.

In the past few years, researchers have attempted to model protein nucleation[4,5] and crystal growth,[6−8] and significant advances have been made in the field of control of crystallization processes introducing new techniques such as the direct nucleation control and statistical control chart based switching.[9,10] Their efforts make it possible to control the shape and size distributions of the protein crystals, however, no significant advance has been made associated with crystal growth initiated by nucleation in the consideration of mass and energy balances. To this end, we include mass and energy balances in order to estimate better the depletion in the protein solute concentration and the drop in the crystallizer temperature due to the heat of fusion by crystallization. Similar to our previous works,[11,12] we assume the solid-on-solid lattice model[13] which will cause the crystal to be very compact by avoiding voids and overhangs while depositing particles onto the crystal lattice, and only monomer units are considered in the attachment events. It is also assumed that the monomer is not aggregated with water, and it is a pure lysozyme molecule.[14] The attachment rate is independent of the local surface microconfiguration, whereas detachment and migration events are highly dependent on it. To account for the dependence of detachment and migration rates on the surface configuration, kinetic Monte Carlo (kMC) simulations are needed to compute the net crystal steady-state growth rate. Kinetic Monte Carlo simulation methods represent a dynamic interpretation of the Master equation[15,16] and have been widely used to simulate dynamic molecular processes.[17−28] To implement our kMC methodology in the consideration of the entire lattice sites and account for mass and energy balances for the continuous phase, we extend the methodology of reference 29 to the rate equations originally developed by Durbin and Feher.[13] The reader may also refer to our previous works for the details of the methodology for single crystal growth[11] and for a set of crystals nucleated at different times,[12] respectively.

Our main contribution is a quantitative comparison of the performance of a novel model predictive control (MPC) strategy to that of two other conventional operating polices: constant

temperature control (CTC) and constant supersaturation control (CSC). The CTC policy attempts to maximize the size of protein crystals by driving the crystallizer conditions from the labile zone to the metastable zone while the CSC policy attempts to maximize the size of protein crystals by maintaining the crystallizer conditions in the metastable zone. Under CTC, the depletion in the solute concentration throughout the entire batch process immediately leads to the drop in the supersaturation level since temperature is constant, and thus less nucleation and lower growth rates are observed. Under CSC, in order to maintain a constant supersaturation level, the depletion in the solute concentration results in the decline in the temperature. Since there is a lower limit on the temperature, if the solute concentration drops too much, a method of simply lowering temperature cannot maintain the supersaturation level to a desired value (i.e., the controller is not robust). This problem steers us to design a new model predictive controller in order to provide optimality as well as robustness and constraint handling in the batch crystallization process.[30,31]

In the next section, we will summarize the rate equations used for our kMC simulation method to describe the surface kinetics of crystallization. Next, we will present derivation of mass and energy balances, and an MPC is designed to achieve the objective of driving the expected growth rate ratio and crystal size to desired set-point values. After that we will show a comparison of the MPC with the conventional operating strategies of constant temperature control (CTC) and constant supersaturation control (CSC), to evaluate the performance of the MPC. Finally, we will finish with a short conclusion and a few ideas for future research.

■ MODELING AND SIMULATION

As said previously, we will use kMC simulations in order to model protein crystal growth. We assume the solid-on-solid lattice model which will cause the growing protein crystal to be very compact by avoiding voids and overhangs. For this work we will focus on square lattice models of length and width $N = 30$ sites with periodic boundary conditions. Previous work[32] demonstrates that no finite size effects were found among systems of sizes $N = 30$, $N = 60$, and $N = 120$ sites. The rate equations for adsorption, desorption, and migration mechanisms, which are similar to those of Durbin and Feher,[13] are introduced in the following subsection. Then, by generating random numbers, each event of our kMC simulation is chosen and executed based on the normalized rates of the three microscopic phenomena. For further details including derivation of the rate equations, update of each lattice height, execution of events, and other details, we refer the reader to our earlier work.[11,12]

Surface Kinetics. As noted in the previous work by Nayhouse et al.[11] and Kwon et al.,[12] the following rate expressions on the crystal surface follow that of Durbin and Feher,[13] which was further developed by Ke et al.[32] for migration events. Every lattice site is considered for attachment where the attachment rate is independent of the surface microconfiguration and defined as,

$$r_a = K^+(\Delta\mu) = K_0^+ \exp\frac{\Delta\mu}{k_B T} \tag{1}$$

where K_0^+ is the attachment coefficient, k_B is the Boltzmann constant, T is the temperature in Kelvin, and $\Delta\mu = k_B T \ln(c/s)$, where c is the solute concentration and s is the protein solubility,

and this term is the crystallization driving force. The protein solubility is dependent on temperature (°C) and defined for pH = 4.5 and 4% (w/v) NaCl by refs 33 and 34 with the following third-order polynomial:

$$s(T) = 2.88 \times 10^{-4}T^3 - 1.65 \times 10^{-3}T^2$$
$$+ 4.619 \times 10^{-2}T + 6.008 \times 10^{-1} \tag{2}$$

where the solubility computed by eq 2 has an error of 6.8 %.[34] The migration rate is modeled by Ke et al.[32] by introducing an additional term to the desorption rate which causes migration to have a higher rate compared to the desorption rate, and the desorption and migration rates are defined the following way:

$$r_d(i) = K_0^+ \exp\left(\frac{\phi}{k_B T} - i\frac{E_{pb}}{k_B T}\right) \tag{3}$$

$$r_m(i) = K_0^+ \exp\left(\frac{\phi}{k_B T} - i\frac{E_{pb}}{k_B T} + \frac{E_{pb}}{2k_B T}\right) \tag{4}$$

where E_{pb} is the average binding energy per bond, $i \in \{0,1,2,3,4\}$ is the number of bonds, and ϕ is the binding energy per molecule of a fully occupied lattice where the binding energies cannot be evaluated by experiments.[13,35] A set of values of E_{pb} and ϕ for (110) and (101) is determined through open-loop kMC simulations until the difference between the calculated and the experimental growth rates becomes negligible.[13] In contrast to the attachment, the detachment and migration event are dependent on the surface microconfiguration because they take into consideration the total binding energy determined by the number of nearest neighbors as it is shown in eqs 3 and 4. Therefore, the growth rate cannot be computed by simply subtracting the detachment from the attachment rates.

In summary, the nature of lysozyme crystals which only half the molecules on the (101) face have the points of attachment for incoming molecules, whereas every molecule on the (110) face has dangling bonds,[13] is reflected in the present study by accepting 50% of adsorption events on the (101) face, compared to 100% of those on the (110) face in the kMC simulation. On the other hand, desorption events are always accepted and migration events are always accepted as long as there exists at least one available migration site. Specifically, an available migration site implies an adjacent site which is lower in height than the current lattice site where a lysozyme molecule can migrate. If there exist multiple sites available for the migrating molecule, a site is chosen randomly.[36]

Mass Balance. Shi et al.[37] modeled the 3-D crystal growth by multiplying a crystal shape factor to the third moment of the crystal size distribution. However, the crystal shape factor is dependent on the temperature and the solute concentration and is not constant. In this work, the mass balance is evaluated by considering the geometry of lysozyme crystal, and thus we are able to model the shape evolution of lysozyme more precisely by distinguishing the crystal growth into (110) and (101) directions independently. In refs 6 and 13 the growth rates for (110) and (101) faces, G_{110} and G_{101}, are respectively related to h_{110} and h_{101} as follows:

$$G_{101} = 0.45\frac{dh_{101}}{dt} \cong 0.45\frac{\Delta h_{101}}{\Delta t} \Rightarrow \quad \Delta h_{101} = 2.22\Delta t G_{101} \tag{5}$$

$$G_{110} = 0.5\frac{dh_{110}}{dt} \cong 0.5\frac{\Delta h_{110}}{\Delta t} \Rightarrow \quad \Delta h_{110} = 2\Delta t G_{110}$$

(6)

where $\Delta t = 1$ s. We model lysozyme as a rectangular prism, shown in Figure 1, whose bottom is a square with a side of h_{110}
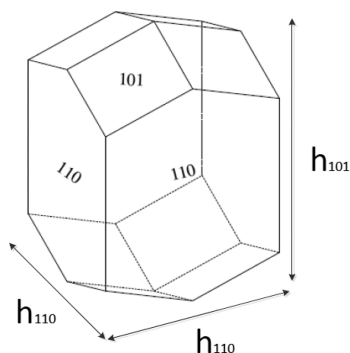

Figure 1. Model of the geometry of lysozyme crystal used in this work.

and a height of h_{101}. From eqs 5 and 6 the crystal size at time $j\Delta t$ can be written as,

$$h_{101}(j\Delta t) = \sum_{k=1}^{j} \Delta h_{101}(k) + h_{101}(0)$$

$$h_{110}(j\Delta t) = \sum_{k=1}^{j} \Delta h_{110}(k) + h_{110}(0)$$

(7)

where $\Delta h_{101}(k) = h_{101}(k\Delta t) - h_{101}((k-1)\Delta t)$ and $\Delta h_{110}(k) = h_{110}(k\Delta t) - h_{110}((k-1)\Delta t)$ for $k = 1, 2, \cdots, j$. Also, the volume of the crystal with a side $h_{110}(t)$ and a height $h_{101}(t)$ at time t, $V_c(t)$, follows as

$$V_c(t) = h_{110}^2(t)h_{101}(t)$$

(8)

and thus the volume change between time $t - \Delta t$ and t, $\Delta V_c(t)$, can be written as,

$$\Delta V_c(t) = V_c(t) - V_c(t - \Delta t)$$

(9)

The amount of the protein solute that is transported from the continuous phase to the crystal at time t can be calculated and takes the following form:

$$V\Delta c(t) = \Delta V_c(t)\rho_c N_c(t)$$

(10)

where V is the volume of the continuous phase (assume change in V is negligible), $N_c(t)$ is the number of crystals in batch at time t, ρ_c is the crystal density, and $\Delta c(t)$ is the change in the protein solute concentration between time $t - \Delta t$ and t where $\Delta c(t) = c(t) - c(t - \Delta t)$. We update the concentration after every Δt (=1.0) seconds. To integrate eq 10 into the model predictive controller below, we estimate the total volume changes of the entire crystals (from i to $N_c(t)$) at time t as follows:

$$\sum_{i=1}^{N_c(t)} \Delta V_{c,i}(t) \approx N_c(t)\langle \Delta V_c(t) \rangle$$

(11)

where the average change in volume over $N_c(t)$ crystals at time t, $\langle \Delta V_c(t) \rangle$, is approximated by

$$\langle \Delta V_c(t) \rangle \approx [\langle h_{110}(t) \rangle^2 \langle h_{101}(t) \rangle - \langle h_{110}(t - \Delta t) \rangle^2$$
$$\langle h_{101}(t - \Delta t) \rangle]$$

(12)

Energy Balance. The energy balance of the batch crystallization process takes the following form given by ref 38.

$$\frac{dT}{dt} = \frac{\rho_c \Delta H_c}{\rho C_p}\frac{d\varepsilon}{dt} - \frac{U_j A_j}{\rho C_p V}(T - T_j)$$

(13)

where $\varepsilon = ((V - V_c)/V) = 1 - (V_c/V)$ is the solids free volume fraction, T is the crystallizer temperature and T_j is the jacket temperature and the manipulated input. The process parameters

Table 1. Parameters for the Batch Crystallizer Model of Equations 12 and 14

ρ_c	crystal density	1400	mg/cm^3
ΔH_c	enthalpy of crystallization	44.5	kJ/kg
ρ	continuous phase solution density	1045	mg/cm^3
C_p	specific heat capacity	4.13	kJ/(K kg)
V	crystallizer volume	5	L
A_j	surface area of heat transfer between crystallizer and jacket pipe	0.25	m^2
U_j	heat transfer coefficient between crystallizer and jacket pipe	1800	kJ/(m^2 h K)

used in the kMC simulations are given in Table 1. Taking a derivative of ε with respect to time and using eq 12, it follows that

$$\frac{dT}{dt} = -\frac{\rho_c \Delta H_c}{\rho C_p}\left(\frac{1}{V}\sum_{i=1}^{N_c(t)} \Delta V_{c,i}(t)\right) - \frac{U_j A_j}{\rho C_p V}(T - T_j)$$

(14)

since

$$\frac{d\varepsilon}{dt} = \frac{d}{dt}\left(1 - \frac{V_c}{V}\right) = -\frac{1}{V}\frac{dV_c}{dt} \cong -\frac{1}{V}\left(\sum_{i=1}^{N_c(t)} \Delta V_{c,i}(t)\right)$$

(15)

Open-Loop Simulation Results. There are many simulation conditions that affect the crystal growth and the nucleation including temperature, pH, salt, and protein solute concentrations. Specifically, the supersaturation σ is defined as $\sigma = \ln(c/s)$ where c (mg/mL) is the protein solute concentration and s (mg/mL) is the solubility which is determined by eq 2 in terms of temperature (°C) at pH = 4.5 and 4% (w/v) NaCl. In Figures 2 and 3, crystal growth has been modeled at supersaturation, $2.1 \lesssim \sigma \lesssim 3.95$, where $c = 45$ (mg/mL). Through a procedure proposed in the previous work of our group,[11,12] our simulation results have been properly calibrated with the experimental result in Figure 2 and a set of parameters was appropriately chosen to verify the crossover behavior of the growth rates of the (110) and (101) faces. The estimated simulation results at 4.0% NaCl are plotted against the experimental results at 3.5% and 5.0% NaCl from ref 6 in Figure 2. Additionally, Figure 3 shows the results of our kMC simulations from Figure 2 along with error bars representing two standard deviations of 10 kMC simulations at each point. The parameters for the kMC simulations used in this work are listed in Table 2.

To complete the kMC simulation model, mass and energy balances are developed to estimate the depletion in the protein solute concentration and the drop in the crystallizer temperature due to the heat of fusion by crystallization. The evolution of the solute concentration, the supersaturation, and the temperature in the kMC simulation is shown in Figure 4 for the initial solute concentration at 47 mg/mL and the three different initial temperatures of 6, 10, and 18 °C. It is observed that the solute
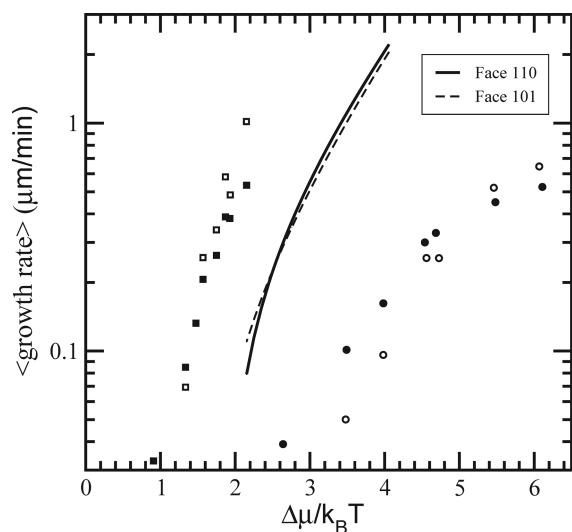
Figure 2. The expected growth rate versus the degree of supersaturation at c = 45 mg/mL and 4% NaCl are shown as the solid (110 face) and the dashed (101 face) lines. The filled square ($\blacksquare$) and open square ($\square$) represent the measured experimental data for 101 and 110 faces with 5% NaCl; ($\bullet$, $\circ$) the measured experimental data with 3.5% NaCl; extracted from ref 6 at pH = 4.6.
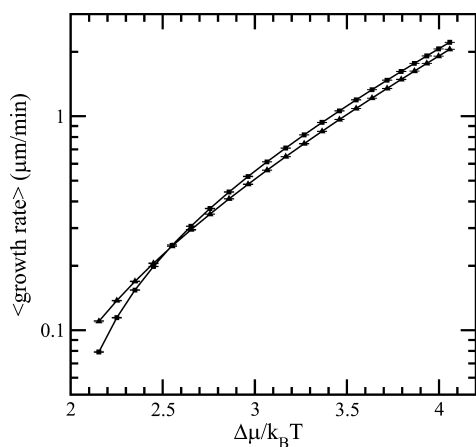

Figure 3. The expected growth rate versus the degree of the supersaturation at c = 45 mg/mL for 110 ($\blacksquare$) and 101 faces ($\blacktriangle$), respectively. Error bars represent two standard deviations of 10 simulations for each point.

Table 2. Parameters for Face (110) and (101) at 45 mg/mL NaCl and pH = 4.5. Additionally, K_0^+ = 0.211 s^{-1}

face	ϕ/k_B	E_{pb}/k_B
(110)	1077.26	227.10
(101)	800.66	241.65

concentration decreases more rapidly as the batch process proceeds further because, for an equivalent growth, larger crystals require more solute to grow compared to small crystals. The temperature of the crystallizer is also affected by the crystallization due to the heat of fusion in the crystallization process. It is verified from Figure 4 that at a low initial temperature level (i.e., low solubility and thus a high super-saturation level) crystals grow faster and it results in a significant drop in the supersaturation level because of the loss in the protein solute for the continuous phase. We note that the drops in the

temperature and the supersaturation depend on the length of batch process time and the size of the crystallizer.

First of all, kMC simulations are run under existing control strategies, a constant temperature control (CTC) strategy and a constant supersaturation control (CSC) strategy; please see below for results. Under CTC, the depletion in the solute concentration throughout the entire batch process immediately leads to the drop in the supersaturation level since temperature is constant, and thus less nucleation and lower growth rates are observed. Under CSC, in order to maintain a constant supersaturation level, the depletion in the solute concentration results in the decline in the temperature. Since there is a lower limit on the temperature, if the solute concentration drops too much, a method of simply lowering temperature cannot maintain the supersaturation level to a desired value (i.e., the controller is not robust). This problem steers us to design a new controller, which is presented in the next section in order to provide optimality as well as robustness in batch process. Specifically, the crystal shape distribution of the final crystals can be driven to a desired range by the controller design described in the following section.

■ MODEL PREDICTIVE CONTROL OF CRYSTAL SIZE AND SHAPE

In the kMC simulations, crystal nucleation and growth are considered alongside mass and energy balances via molecular attachment, detachment, and migration events. Since the role of mass and energy balances becomes significant as crystal size increases, the balance equations have been considered in the controller design. The nucleation and the crystal growth rates have been manipulated by changing the temperature for a given concentration. In Table 2, parameters of crystal growth conditions (e.g., E_{pb}/k_B and ϕ/k_B) for kMC simulations are chosen to show the experimentally observed crossover behavior in the crystal growth rates between the (110) and (101) faces.[6,13] In addition to the nonlinear models described in Figure 5 panels a−c, mass and energy balances introduced play a key role in describing the behavior of the system dynamics in the consideration of the depletion in the solute concentration and the heat removal in the crystallizer due to crystallization process. Then, based on these equations, a model predictive controller is designed to produce the crystals with the desired shape and size through the manipulation of the jacket temperature. MPC resolves the drawbacks of the classical control schemes like proportional (P) control as it explicitly takes into consideration input/state constraints, optimality issues, and nature of non-linearity in the nonlinear growth rate. A dynamic open-loop optimization method may be used. However, open-loop optimizations are not robust with respect to model uncertainty in the protein concentration and the batch crystallizer.

The Population Balance Equation for Protein Crystallization. A population balance equation (PBE), shown in eq 18 below, can describe the evolution of the particle growth in a batch protein crystallization with respect to crystal size and shape.

$$\frac{\partial n(h_{110}, h_{101}, t)}{\partial t} + G_{110}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}}$$
$$+ G_{101}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{101}}$$
$$= B(T, c)\delta(h_{110}, h_{101}) \qquad (16)$$

(a) concentration
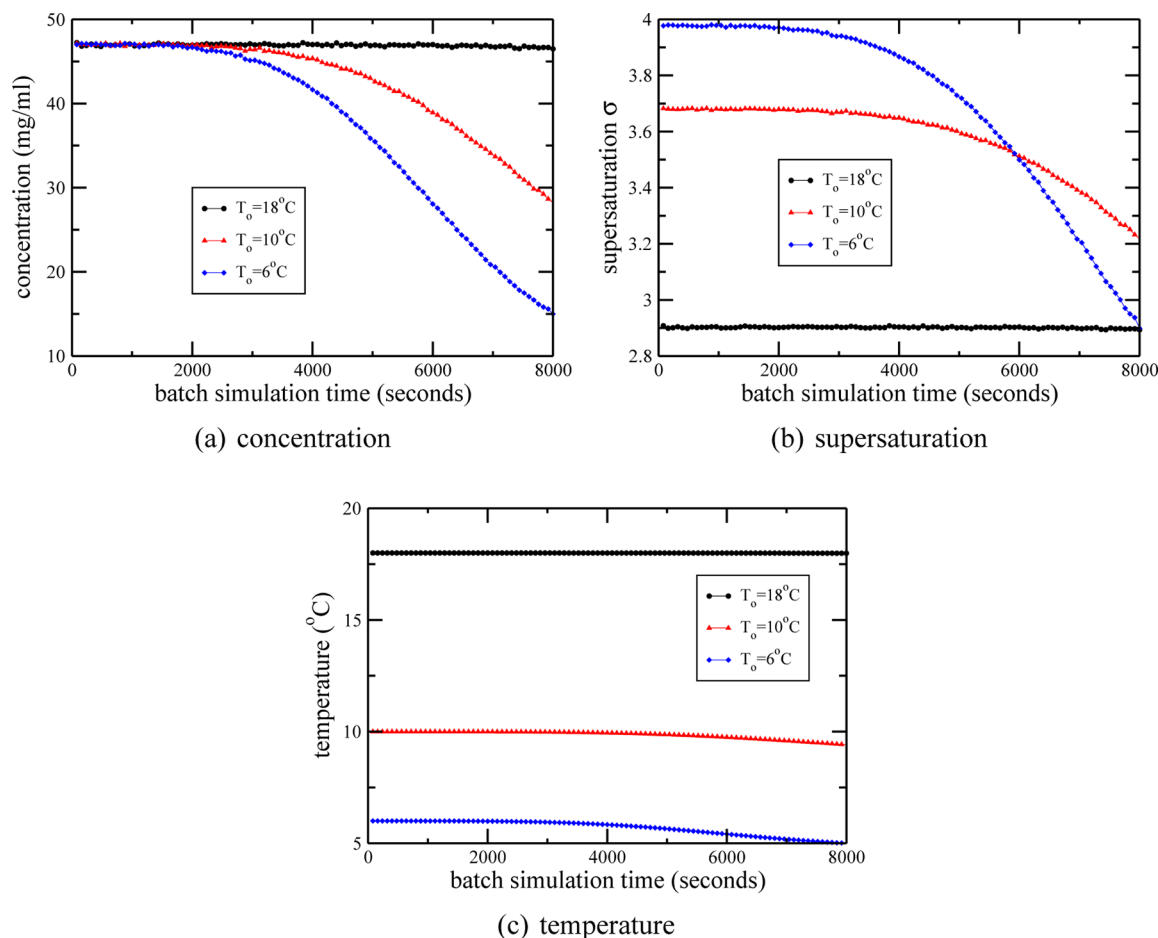

(b) supersaturation

(c) temperature

Figure 4. Open-loop simulation results of solute concentration, supersaturation, and temperature for tetragonal lysozyme protein crystals at pH = 4.5 under the consideration of the mass and the energy balances for the continuous phase. The data from the open-loop simulation shows the depletion in the protein solute concentration and the drop in the temperature due to crystallization processes. The initial protein concentration is 47 mg/mL and the different initial temperature values (6, 10, and 18 °C) are used to verify their effect on the crystal growth.

To describe the behavior of the crystal size and shape distributions for the crystals nucleated at different times in a crystallization process, it is necessary to know the nucleation rate. In kMC simulation the characteristic crystal lengths are simply h_{110} and h_{101} as they are in Figure 2 and $n(h_{110}, h_{101}, t)$ is the number of crystals of heights h_{110} and h_{101} for (110) and (101) faces, respectively, at time t. The nucleation rate is denoted as $B(T,c)\delta(h_{110}, h_{101})$ and is a function of temperature, T, and protein solute concentration, c, which acts only at $h_{110} = h_{101} = 0$. Owing to the dependence of detachment and migration rates on the surface configuration, Nayhouse et al.[11] suggested kMC simulations to compute the net, steady-state, growth rate as a function of temperature and protein concentration in the continuous phase. Therefore, the growth rates for (110) and (101) faces, $G_{110}(T,c)$ and $G_{101}(T,c)$, are expressed as a function of temperature, T, and protein solute concentration, c, respectively. Then, we define a growth rate ratio, $\alpha(T,c) = h_{110}/h_{101} = G_{110}(T,c)/G_{101}(T,c)$, where the growth rate ratio is equal to the aspect size ratio at the steady-state and the growth rate ratio is obtained from kMC simulation. Then, eq 16 can be written as follows,

$$\frac{\partial n(h_{110}, h_{101}, t)}{\partial t} + G_{110}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}}$$

$$+ G_{101}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}}\frac{\partial h_{110}}{\partial h_{101}}$$

$$= B(T, c)\delta(h_{110}, h_{101}) \tag{17}$$

where $\partial h_{110}/\partial h_{101} = \alpha(T,c)$, and substituting this into eq 17, we obtain,

$$\frac{\partial n(h_{110}, h_{101}, t)}{\partial t} + (G_{110}(T, c) + \alpha(T, c)G_{101}(T, c))$$

$$\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}} = B(T, c)\delta(h_{110}, h_{101}) \tag{18}$$

In the present work, we assume that crystals are nucleated at negligibly small size (i.e., $h_{110} = h_{101} = 0$), and the number of nuclei newly formed at time t is denoted as $n(0,0,t)$. This assumption can be justified by the following reasons. Initially, an HEW lysozyme nucleus is formed through the aggregation of 3 to 4 lysozyme molecules, and its size is relatively infinitesimal compared to the final crystal size which is in the dimension of several hundreds micrometers.[39] Additionally, a nuclei cannot be detected until its size reaches the resolution limit, ~0.5 μm.[40] Because the $B(T,c)\delta(h_{110},h_{101})$ acts only at $h_{110} = h_{101} = 0$, and a

(a) Growth rate ratio

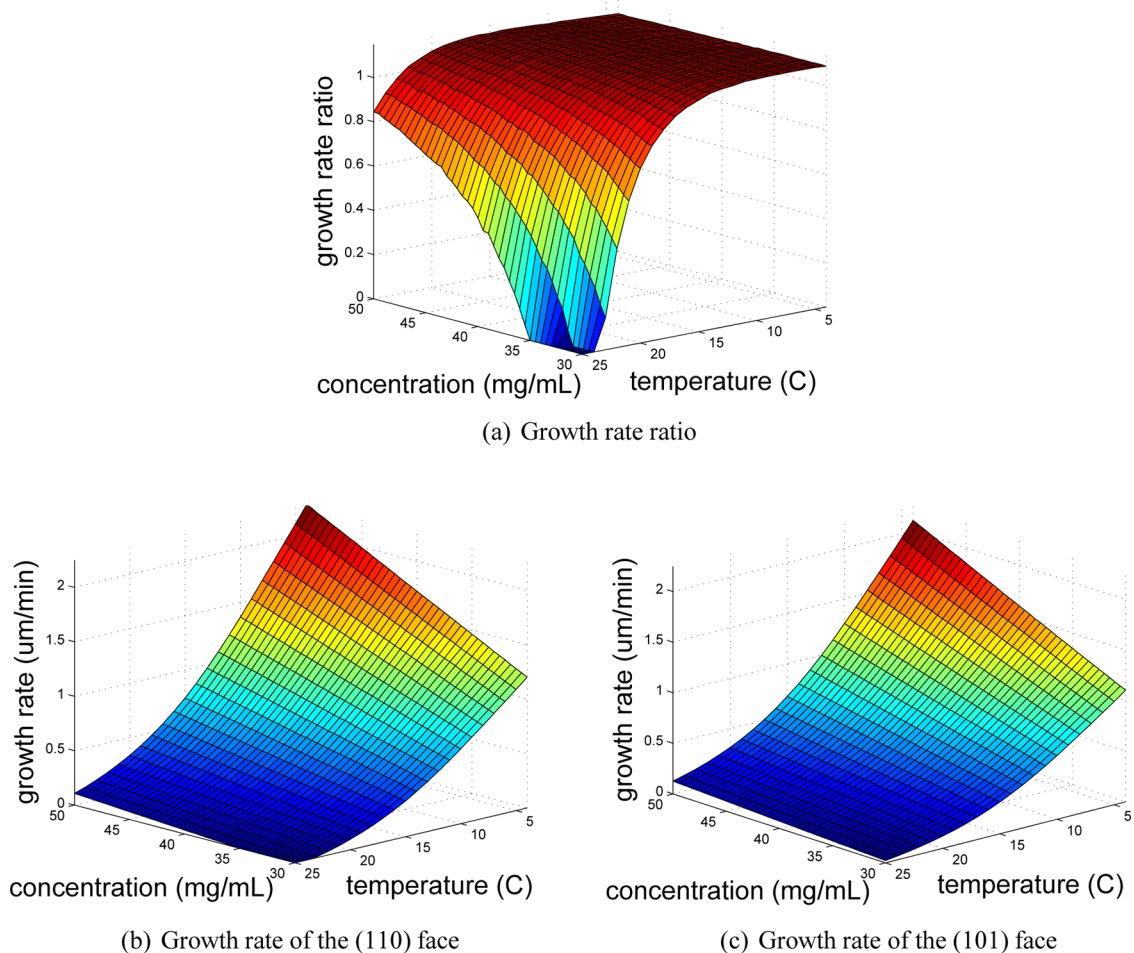(b) Growth rate of the (110) face

(c) Growth rate of the (101) face

Figure 5. Plots of the growth rate data obtained for the (110) face, the (101) face, and the growth rate ratio between (110) and (101) faces for tetragonal lysozyme protein crystals at pH = 4.5. The data from the open-loop kMC simulation are plotted to demonstrate the effect of temperature and concentration variations on growth rates. Protein concentration and temperature range from 30 to 50 mg/mL and 4 to 25 °C, respectively.

nuclei size on (110) and (101) direction is uniformly negligible, for the simplicity of calculations, we can assume that $B(T,c)\delta(h_{110}, h_{101}) \cong B(T,C)\delta(h_{110})$. In other words, the crystal size in the direction of (110) and (101) faces are in the same order of magnitude. As is pointed out by Miller,[41] eq 18 can be written replacing the term with an appropriate boundary condition. Integrating eq 18 over h_{110} from 0^- to 0^+ gives,

$$(G_{110}(T, c) + \alpha(T, c)G_{101}(T, c))$$
$$\times \int_{0^-}^{0^+} \frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}} \, dh_{110}$$
$$= B(T, c) \tag{19}$$

because,

$$\int_{0^-}^{0^+} B(T, c)\delta(h_{110}) \, dh_{110} = B(T, c)$$

and,

$$\int_{0^-}^{0^+} \frac{\partial n(h_{110}, h_{101}, t)}{\partial t} \, dh_{110} = 0$$

It also follows that $n(h_{110}, h_{101}, t) = 0$ at $h_{110} = h_{101} = 0^-$ and $n(h_{110}, h_{101}, t) = n(0,0,t)$ at $h_{110} = h_{101} = 0^+$, assuming all nuclei form with size $h_{110} = h_{101} = 0$. Then eq 19 can be reduced to the following:

$$(G_{110}(T, c) + \alpha(T, c)G_{101}(T, c))n(0, 0, t) = B(T, c) \tag{20}$$

Therefore the desired boundary condition is

$$n(0, 0, t) = \frac{B(T, c)}{(G_{110}(T, c) + \alpha(T, c)G_{101}(T, c))}$$

$$\text{at } h_{110} = h_{101} = 0 \tag{21}$$

and with this boundary condition the resulting population balance of eq 16 has the following form:

$$\frac{\partial n(h_{110}, h_{101}, t)}{\partial t} + G_{110}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{110}}$$
$$+ G_{101}(T, c)\frac{\partial n(h_{110}, h_{101}, t)}{\partial h_{101}} = 0$$

$$\alpha(T, c) = f_G(T, c)$$

$$G_{110}(T, c) = f_{110}(T, c, t)$$

$$G_{101}(T, c) = f_{101}(T, c, t)$$

$$B(T, c) = f_{\text{nucleation}}(\sigma(t))$$

$$n(0, 0, t) = \frac{B(T, c)}{(G_{110}(T, c) + \alpha(T, c)G_{101}(T, c))}$$
$$\text{at } h_{110} = h_{101} = 0 \tag{22}$$

where $G_{110}(T,C)$ and $G_{101}(T,C)$ are the growth rates for (110) and (101) faces, $\alpha(T,c) = G_{110}(T,C)/G_{101}(T,C)$, is the growth rate ratio. The nonlinear equations, f_G, f_{110}, and f_{101}, show their dependencies on temperature, solute concentration, and time, respectively. In this work, it is assumed that the nucleation on a surface is negligible. The number of crystals nucleated at time, t, is obtained from Galkin et al.[42] and the nucleation rate, $f_{\text{nucleation}}(0,0,t)(\sigma)$, at time t (with units $[\text{cm}^{-3}\cdot\text{sec}^{-1}]$), was obtained from ref 42 at pH = 4.5 and 4% (w/v) NaCl:

$$f_{\text{nucleation}}(0, 0, t)(\sigma)$$
$$= \begin{cases} 0.041\sigma + 0.063 & \text{for } \sigma \geq 3.11 \\ 8.0 \times 10^{-8} \exp(4.725\sigma) & \text{for } \sigma < 3.11 \end{cases} \tag{23}$$

We execute multiple kMC simulations alongside mass and energy balances for crystals nucleated at different times which is considered to be comparable to solving eq 22 directly. Theoretically speaking, it requires an infinite number of lattice sites in the kMC simulation to completely regenerate the deterministic PBM described by eq 22. From a practical standpoint, however, a kMC simulation with a finite number of lattice sites is used for the simulation of crystallization process which may lead to a mismatch between the PBM and the kMC simulation. Here we assume that the number of lattice sites being used in the kMC simulation is sufficient that the mismatch is negligible since, as described previously, no size effects were found in the systems with more lattice sites.

Model Predictive Control Formulation. We consider the control of shape and size of crystals nucleated at different times along with mass and energy balances as the batch crystallization process proceeds by using a model predictive control (MPC) design. Minimizing the expected value of the growth rate ratio, $\langle \alpha \rangle = \langle G_{110}/G_{101} \rangle$, is chosen as the control objective. To prevent the crystallizer from obtaining many small crystal fines at the end of the batch run, a desired minimum crystal size is considered in the cost function of the MPC formulation. Among various factors that affect the evolution of the crystal morphology and the growth rate during the crystallization process,[43−47] the jacket temperature is used as the manipulated input, and only the solute concentration measurement is available, assuming all other parameters remain constant for the closed-loop simulations (e.g., pH, NaCl concentration, buffer concentration, etc.).

We note that the proposed modeling and control methods can be extended to the case of multiple manipulated variables. A number of practical considerations including mass and energy

balances of eqs 12 and 14 and additional constraints are considered in the control problem. First, a constraint on the range of the jacket temperature is imposed to ensure that the protein is not damaged where $4\,°C \leq T_j \leq 25\,°C$. Second, there is a constraint on the rate of change of the jacket temperature because of actuator limitations of $2\,°C/\text{min}$. The other constraint limits the number of crystals nucleated during the second half of the batch run to avoid small crystal fines at the end of the batch run. The control action (jacket temperature) at time t is obtained by solving a finite-dimensional optimization problem in a receding horizon fashion. The cost function in the optimal control problem includes a penalty on the deviation of $\langle \alpha \rangle$ from its desired crystal shape. An additional penalty cost is included to account for the negative deviation of the crystal size when its size is less than the desired minimum. In the proposed MPC, crystal growth and nucleation are estimated by using the nonlinear equations, Figure 5, and eq 23, respectively. The proposed MPC formulation is presented as follows:

$$\underset{T_{j,1},\ldots,T_{j,i},\ldots,T_{j,p}}{\text{minimize}} \sum_{i=1}^{p} F_{\langle\alpha\rangle,i} + F_{h_{110},i} + F_{h_{101},i}$$

subject to

$$F_{\langle\alpha\rangle,i} = (\langle\alpha\rangle - \alpha_{\text{set}})^2$$

$$F_{h_j,i} = \begin{cases} \dfrac{h_{j,\min} - \langle h_j(t_i)\rangle}{h_{j,\min}} & \text{for } \langle h_j(t_i)\rangle < h_{j,\min} \\ 0 & \text{for } \langle h_j(t_i)\rangle \geq h_{j,\min} \end{cases}$$

$$G_i = f_r(T, C, t_i)$$

$$G_j(t_i) = f_j(T, C, t_i)$$

$$T_{\min} \leq T_{j,i} \leq T_{\max}$$

$$\left| \frac{T_{j,i+1} - T_{j,i}}{\Delta} \right| \leq R_T$$

$$\frac{dT}{dt} = -\frac{\rho_c \Delta H_c}{\rho C_p}\left(\frac{1}{V} \sum_{i=1}^{N_c(t)} \Delta V_{c,i} \right) - \frac{U_c A_c}{\rho C_p V}(T - T_{j,i})$$

$$\Delta c(t) = \frac{\rho_c}{V} \sum_{i=1}^{N_c(t)} \Delta V_{c,i}$$

$$n(0, 0, t) \leq n_{\text{limit}} \quad \forall\, t \geq t_f/2$$

$$\langle h_j(t_i)\rangle = \frac{\langle h_j(t_{i-1})\rangle n(h_{110}, h_{101}, t_{i-1})}{n(h_{110}, h_{101}, t_i)} + R_j(t_{i-1})\Delta$$

$$n(h_{110}, h_{101}, t_i) = n(h_{110}, h_{101}, t_{i-1}) + n(0, 0, t_{i-1}) \quad \forall\, i$$

$$i = 1, 2, \ldots, p$$

$$j \in \{110, 101\} \tag{24}$$

where t is the current time, t_i, $i = 1, 2, \ldots, p$, is the time of the ith prediction step, $t_i = t + i\Delta$, respectively, t_f is the total time of the batch simulation, $F_{\langle\alpha\rangle,i}$ is the cost function expressing the deviation of $\langle\alpha\rangle$ from its set-point ratio, α_{set}, $F_{h110,i}$ and $F_{h101,i}$ are the cost functions expressing the penalty on the negative

deviation of $\langle h_{110} \rangle$ and $\langle h_{101} \rangle$ from its minimum crystal size, $h_{110,min}$ and $h_{101,min}$, at time t_i, p is the number of prediction steps, $p\Delta$ is the specified prediction horizon, $T_{j,i}$, $i = 1, 2, ..., p$, is the jacket temperature at the ith step, $T_{j,i} = T_j(t + i\Delta)$, respectively, T_{min} and T_{max} are the lower and upper bounds on the jacket temperature, respectively, R_T is the limit on the rate of change of the jacket temperature, n_{limit} limits the number of crystals nucleated during the latter half of the simulation time. The number of crystals, $n(h_{110}, h_{101}, t_i)$, and average height of the crystal face j, $\langle h_j(t_i) \rangle$, at time t_i are updated at every sampling time through the recursive equations (cf. eq 24), respectively. The set of optimal jacket temperature, $(T_{j,1}, T_{j,2}, ..., T_{j,p})$, is obtained by solving the multivariable optimization problem of eq 24, and only the first value of the optimal jacket temperature trajectory, $T_{j,1}$, is applied to the protein crystallization process until the next sampling time. Then, a new measurement of protein concentration in the continuous phase is received from the kMC simulation, and the MPC problem of eq 24 is resolved for the computation of the next optimal input trajectory. In a previous work,[37] empirical expressions were used to simulate the crystal growth and nucleation. In the present work, however, the kMC simulations are executed based on the rate equations described previously to simulate the crystallization process to a higher degree of precision. Furthermore, the uncertainty in the system and the model mismatch will be taken into account in the protein concentration variations. For further results including robust-control of crystallization systems and model predictive control, the reader may refer to ref 48 and ref 30, respectively.

■ BATCH CRYSTALLIZATION UNDER CLOSED-LOOP OPERATION

In this section, the proposed model predictive controller of eq 24 is solved via a local constrained minimization algorithm using the nonlinear algebraic models described previously (cf. Figure 5) which show the solute concentration and temperature dependencies of the crystal growth rate and growth rate ratio, respectively. At each sampling time (1 s), the optimal jacket temperature, obtained by solving the optimization problem of eq 24, is applied to the closed-loop system until the next sampling time.

The solute concentration randomly fluctuates following the Gaussian distribution given by eq 25 below to simulate the uncertainty in the system at pH 4.5 and 4.0% NaCl, i.e.,

$$\langle C(t) \rangle = C_n, \quad \langle C(t)C(t') \rangle = \sigma_n^2 C_n^2 \tag{25}$$

where C_n is the nominal concentration of the system and σ_n^2 shows how far a set of measured concentrations deviates from its nominal value. We also note that the concentration variation affects the attachment rate [cf. eq 1]. For all closed-loop simulations, the nominal concentration is 44 mg/mL, and the deviation, σ_n, is equal to 1.5%. The maximum rate of change of the jacket temperature is 2 °C/min. The volume of the crystallizer is 5.0 L. We note that we have taken into account the following heuristic: In the beginning, the crystallizer operates in the labile zone where the supersaturation level is so high that both the nucleation and the crystal growth occur. Then the crystallizer operates in the metastable region which is a relatively high supersaturation region where nucleation rarely occurs, but the crystal growth still does. This will help prevent the small crystal fines appearing at the end of the batch run, and this heuristic is taken into consideration in this work as one of the constraints restricting the number of crystals nucleated during

the latter half of the simulation time, $n_{limit} = 500$. Since the MPC formulation uses steady-state growth rates (cf. Figure 5b,c) which implies that it is a slowly varying process, the number of prediction steps is set to be $p = 1$. The time interval between the two sampling times is 1 s. The prediction horizon of each step is fixed at $p\Delta = 1$ s. For the purpose of simulation, the solute concentration is set to be fluctuating every 1 s with the nominal value $C_n(t)$ at time t. The computational time that is used to solve the optimization problem with the current available computing power is negligible with respect to the sampling time interval. The closed-loop simulation duration $t_f = 8000$ s.

In the closed-loop simulations the control objective is to drive the expected growth rate ratio to the desired set-point values, $\langle \alpha \rangle = 0.85$ and $\langle \alpha \rangle = 1.11$. We chose these two values to represent two different crystal morphologies available with lysozyme crystals. For the former set-point value, the protein crystal shape is slightly elongated along the (101) direction while it is more equidimensional for the latter case. Thus, the cost function of this problem contains a penalty on the deviation of the expected growth rate ratio from the desired shape.

We compare the performance of the proposed MPC to that of two other conventional control strategies, constant temperature control (CTC) and constant supersaturation control (CSC). Compared to MPC, the crystal shape distribution as well as the solute concentration, the temperature, and the supersaturation propagates in a very different manner under CTC and CSC. Under CTC, the crystallizer temperature is maintained constant which results in a constant solubility during the batch run, and thus, the depletion in the solute concentration eventually leads to the decrease of supersaturation. Under CSC, however, the controller tries to keep a constant supersaturation throughout the batch run. This is done by constantly decreasing the solute concentration, thus maintaining the supersaturation level nearly constant throughout the batch. Under CSC, therefore, the growth rate which only depends on the supersaturation itself remains constant during the batch run. The nucleation rate also stays constant under CSC since it is only dependent on the supersaturation. In the case of MPC and CTC, however, a biased nucleation occurs (e.g., 50% of crystals nucleate in the first 10% of the entire batch simulation time) throughout the batch run due to the change in the supersaturation because of the depletion in the solute concentration or the drop in the crystallizer temperature. We note that if the temperature reaches its optimal state relatively later in the batch run, the biased nucleation could result in a broader crystal shape distribution from a desired set-point value compared to that of the operation under CSC.

For the lower growth rate ratio set-point value, $\langle \alpha \rangle = 0.85$, the results of the closed-loop simulations are shown in Figures 6, 7, and 8 with respect to the crystallizer temperature, the solute concentration, the supersaturation, the nucleation time distribution, and the crystal shape distribution at the end of the batch run. Specifically, Figure 6 shows results for batch runs under three different control strategies (i.e., MPC, CTC at three different temperature values, and CSC at three different supersaturation values). Note that we chose 3 values for the initial temperature for the simulations under CTC at $T_o = 13$ °C, $T_o = 18$ °C, and $T_o = 23.4$ °C which are two extremes, and one in the middle of the temperature trajectory computed by the MPC executed at $c = 44$ mg/mL and $T_o = 13$ °C. The three values for the supersaturation were chosen in the same manner for the simulations under CSC at $\sigma = 2.34$, $\sigma = 2.84$, and $\sigma = 3.41$, respectively. It is also noted that two different sets of the solute concentration and the temperature may result in the same supersaturation σ, and it is
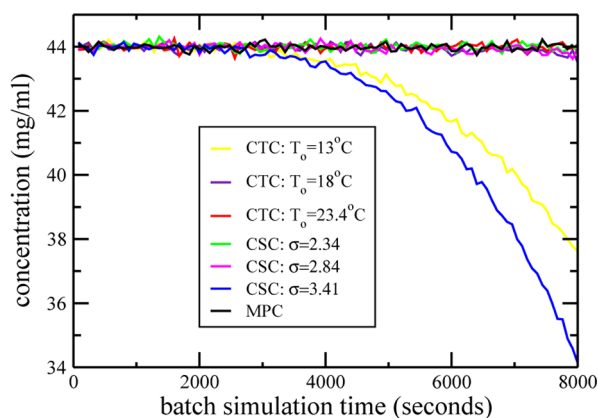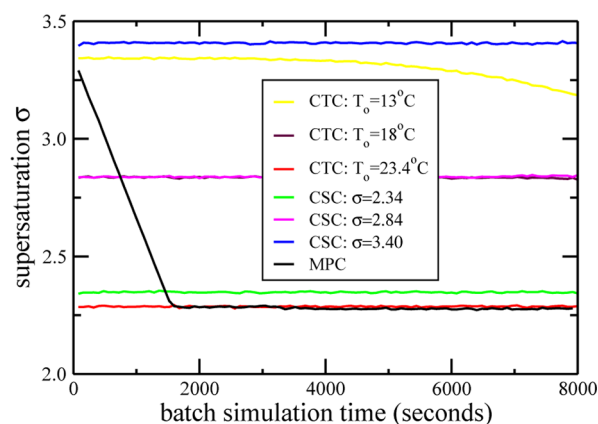
(a) concentration vs time
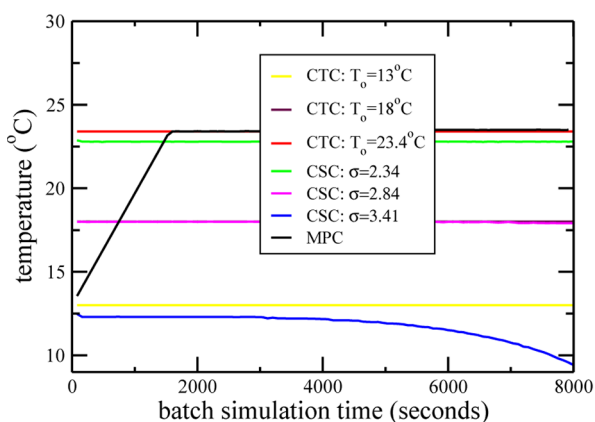
(b) temperature vs time

(c) supersaturation σ vs time

Figure 6. The propagation of temperature, concentration, and supersaturation with time during the batch run under closed-loop operation at different initial temperature values and supersaturation levels alongside MPC aiming at growth rate ratio set-point value, $\langle \alpha \rangle = 0.85$.

the main driving force for the crystal growth and nucleation. Therefore the evolution of supersaturation can be used with nucleation time distribution, Figure 7, to explain the behavior of
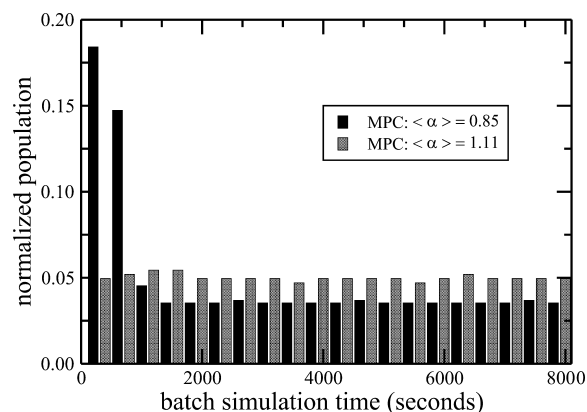

Figure 7. Profiles of nucleated crystals with time during the batch run under closed-loop operation under MPC aiming for different growth rate ratio set-point values, $\langle \alpha \rangle = 0.85$ and $\langle \alpha \rangle = 1.11$, respectively. It is noted that the nucleation time distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars, for each different set of growth rate ratio, will add up to 1.
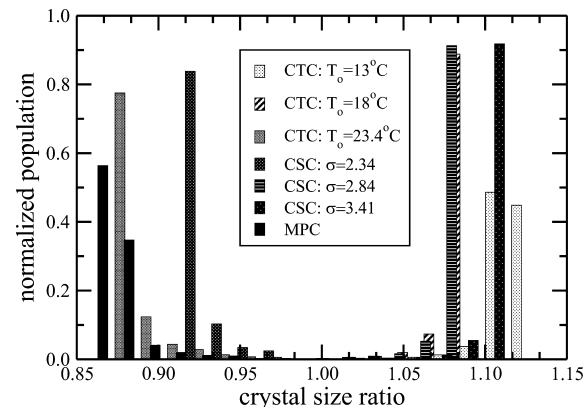

Figure 8. The final crystal shape distribution at the end of the batch simulation under different closed-loop operations under CTC and CSC at different initial temperature values and supersaturation levels for CTC and CSC, respectively, and MPC for the growth rate ratio set-point value, $\langle \alpha \rangle = 0.85$. It is noted that the crystal shape distribution is a dimensionless variable and is normalized over the entire population so that summing over all histograms will add up to 1 for each control strategy.

the crystal shape distribution under MPC for different set-point values.

It has been demonstrated in our recent work[12] that the final crystal shape distribution is very narrow when the initial temperature is close to the optimal temperature. Additionally, depending on the desired crystal morphology and the initial temperature of the crystallizer, it takes a different amount of time until the temperature reaches the desired set-point value. Specifically in Figure 8, for the lower desired crystal growth rate ratio, $\langle \alpha \rangle = 0.85$, the initial temperature of the crystallizer was chosen to be $T_o = 13$ °C although it is not very close to the optimal temperature ~23.4 °C. Therefore, starting from a high initial temperature is encouraging, because the system reaches its optimal temperature faster which enables the crystals to uniformly nucleate along the batch and they undergo an optimal temperature profile from the beginning. Note that the proposed MPC with a relatively low initial temperature still outperforms other conventional polices as it is seen in Figure 8, and the

performance of the MPC can be further improved by using a higher initial temperature. Figure 6a displays that the solute concentration has depleted significantly for the simulations under the CSC at $\sigma = 3.41$ and the CTC at $T_o = 13\ °C$ which both result in high supersaturation levels, and hence, high growth and nucleation rates. On the other hand, the proposed MPC, CTC at the relatively high temperature, $T_o = 18\ °C$ and $T_o = 23.4\ °C$, and the CSC at the low supersaturation, $\sigma = 2.34$ and $\sigma = 2.84$, show no significant changes in their solute concentration because of a low growth rate resulting from a low supersaturation level. We note that in Figure 6, the solute concentration, the supersaturation, and the temperature with time under CTC at $T_o = 18\ °C$ are very similar to those of CSC at $\sigma = 2.84$. The resulting crystal size and shape distributions are very similar as it is shown in Figure 8 and Table 3 which implies that CTC and CSC

Table 3. Comparison between the Simulation Results for the Crystal Size in (110) Direction under Three Different Control Strategies for the Desired Growth Rate Ratio of $\langle \alpha \rangle = 0.85$

| control strategy	$\langle h_{110} \rangle$	r_{10}	r_{50}	r_{90}	span
MPC	11.83	3.15	12.56	18.82	1.25
CTC: $T = 13\ °C$	58.41	11.29	57.79	106.68	1.65
CTC: $T = 18\ °C$	29.78	6.52	29.73	52.96	1.56
CTC: $T = 23.4\ °C$	3.39	0.92	3.39	5.87	1.46
CSC: $\sigma = 2.34$	3.96	0.95	3.93	6.91	1.52
CSC: $\sigma = 2.84$	29.75	6.50	29.63	52.79	1.56
CSC: $\sigma = 3.41$	67.98	14.19	67.94	121.79	1.58

policies result in a similar performance when their growth and nucleation rates are low and their initial conditions are identical. Owing to the negligible depletion in the solute concentration, the black solid line in Figure 6b shows that the MPC computes a jacket temperature that is first monotonically increasing to the optimal value and then it is staying constant. Usually for a low growth rate ratio, the optimal temperature is high where the system is usually very sensitive to the solute concentration changes as it is shown in Figure 5. Specifically, the growth rate ratio declines drastically at high temperature, where the desired low growth rate ratio is available, in response to the small variation in the concentration. In Figure 7, for $\langle \alpha \rangle = 0.85$, the optimal temperature is ~23.4 °C and high initial supersaturation levels result in the nucleation of 34% of the total crystals within the first 500 s of the batch run. Although the MPC in this case, owing to the favored nucleation in the beginning, results in a similar crystal shape distribution compared to that of CTC at $T_o = 23.4\ °C$ where the constant optimal temperature is maintained for the crystallizer over the entire batch. The final crystal shape distribution for the MPC in Figure 8 can be very narrow and even closer to the desired values by adjusting the initial temperature closer to the optimal temperature, from $T_o = 13\ °C$ to $T_o = 23.4\ °C$. Evaluation of the sensitivity of the controller design with respect to the initial temperature of the crystallizer has been presented in our recent work.[12]

In Table 3, the characteristics of crystal size along the (110) direction of the final crystals at the end of the batch run are compared under different control strategies, including the expected crystal size along (110) direction, $\langle h_{110} \rangle$, and r_{10}, r_{50}, r_{90}, which are the 10%, 50%, and 90% population fractions of the crystal size distribution for h_{110}, respectively, representing the percentage of population at crystal size less than that value. Only the crystal size for the (110) direction, $\langle h_{110} \rangle$, is included since h_{110} and h_{101} are within the same order of magnitude. This table

also includes the span which is defined as $(r_{90} - r_{10})/r_{50}$, and it is a widely used characteristic in the pharmaceutical industries. Comparing the results of the five control strategies listed in Table 3, it is clear that the MPC is able to increase the crystal size and achieve a low polydispersity while it also drives the expected crystal shape, $\langle \alpha \rangle$, to the desired value as it is seen in Figure 8. Although the CTC at $T_o = 23.4\ °C$ results in a similar $\langle \alpha \rangle$ to that of MPC, it leads to much smaller crystal size with a high polydispersity, where neither of them is desirable. The very low span value for the MPC indicates a narrow crystal size distribution (a low polydispersity) which is obtained by properly dealing with the biased nucleation rate described previously. Again, the controller performance can be improved if we choose an initial temperature as close as possible to the optimal temperature.

In contrast to $\langle \alpha \rangle = 0.85$, the case of the higher desired crystal growth rate ratio, $\langle \alpha \rangle = 1.11$, has shown that changes in the solute concentration are no longer negligible which can be attributed to the high growth and nucleation rates. Specifically, the solid black line in Figure 9 shows that the crystallizer temperature manipulated by the proposed MPC results in the supersaturation first increasing due to a drastic drop in the crystallizer temperature, then decreasing almost to the lowest level, which is then followed by a series of rise-and-fall until the end of the batch run. Since the control objective is to obtain a very narrow final crystal shape distribution centering around a desired set-point value, the result can be understood as follows: a drastic initial rise in the supersaturation (from $\sigma = 3.4$ to $\sigma = 3.87$) which does not significantly increase the number of crystals nucleated because a small increase in the supersaturation level does not result in a substantial increase in the nucleation rate for $\sigma > 3.11$ as it is shown in eq 23; then the supersaturation drops to a low level in order to minimize the nucleation rate as the growth regime is changing from the labile to the metastable zone. Thereafter, the series of rise-and-fall is attributed to the optimal jacket temperature trajectory computed from the proposed MPC which drives the system to the vicinity of the optimal growth conditions for a desired growth rate ratio while dealing with the solute concentration reduction from 44 to 35 mg/mL as seen in Figure 9a−c. Note that a high supersaturation level, which is necessary for the desired equidimensional shape, can result in a significant solute concentration drop due to a higher crystal nucleation rate as well as due to very fast crystal growth. The crystal growth rate ratio, however, is dependent on supersaturation which is the ratio between the solute concentration and the solubility. Thus, a desired supersaturation level can be maintained by appropriately decreasing the solubility, which is a function of the crystallizer temperature. In Figure 9b, the computed trajectory for the crystallizer temperature by the proposed MPC is similar to that of the standard CSC; however, it is different in a sense that the MPC deals with the nuclei formation in order not to have small crystal fines at the end of the batch process. In Figure 7, the crystallizer was run with a temperature which is relatively close to the optimal temperature of ~11 °C and, as a result, crystals nucleate uniformly throughout the batch run. Because of the insensitivity of the system to variations in the temperature and solute concentration at a high desired crystal growth rate ratio, the crystal shape distribution under MPC is similar to that of the MPC at the low growth rate ratio, $\langle \alpha \rangle = 0.85$, as it is seen in Figure 10.

As compared to the $\langle \alpha \rangle = 0.85$, in Table 4 a high span (a high polydispersity) but a narrow and desired crystal shape distribution is resulting from a nearly uniform crystal nucleation
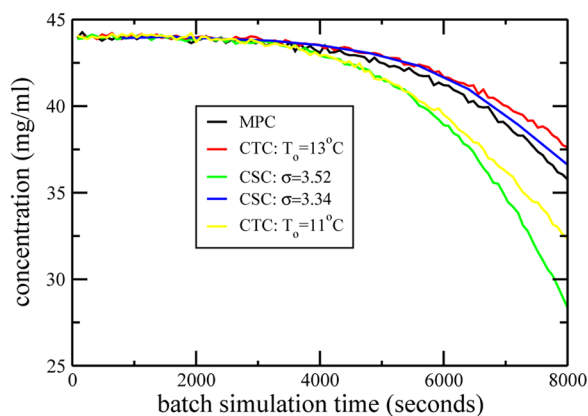
(a) concentration vs time
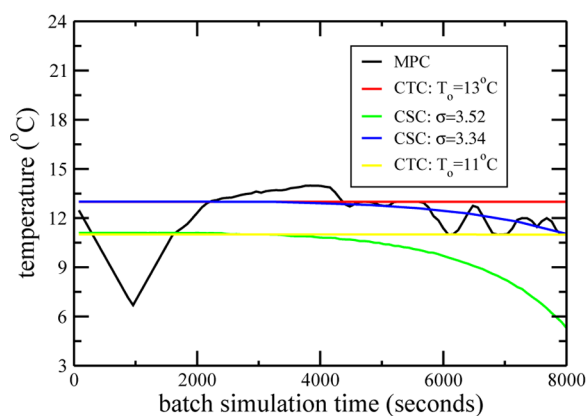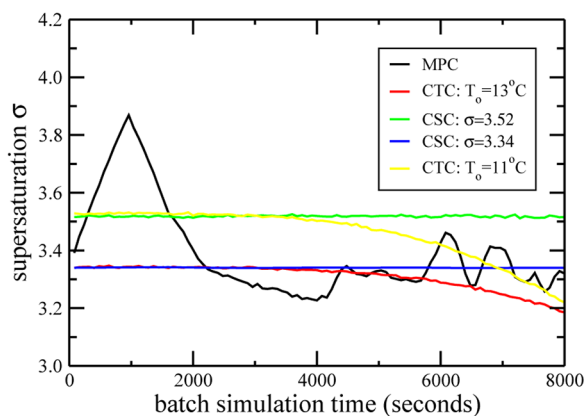

(b) temperature vs time

(c) supersaturation σ vs time

Figure 9. The propagation of temperature, concentration, and supersaturation with time during the batch run under closed-loop operation at different initial temperature values and supersaturation levels alongside MPC aiming at growth rate ratio set-point value $\langle \alpha \rangle = 1.11$.

rate under MPC compared to the simulations under CTC and CSC. In addition, the MPC successfully deals with the significant concentration drop to maintain the desired crystal shape, $\langle \alpha \rangle = 1.11$. Although the CTC and CSC result in similar shape distributions to that of MPC, their performances are not robust
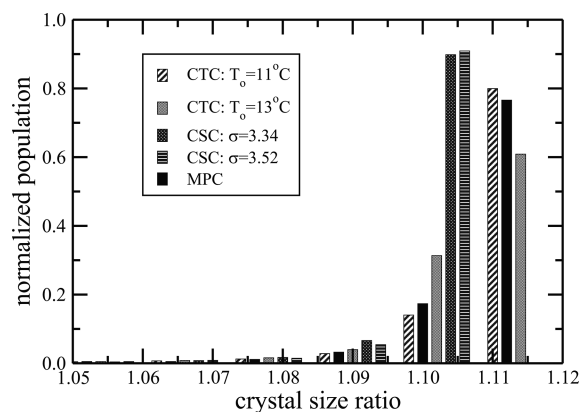

Figure 10. The final crystal shape distribution at the end of the batch simulation under different closed-loop operations under CTC and CSC at different initial temperature values and supersaturation levels for CTC and CSC, respectively, and MPC for the growth rate ratio set-point value $\langle \alpha \rangle = 1.11$. It is noted that the crystal shape distribution is a dimensionless variable and is normalized over the entire population so that summing over all histograms will add up to 1 for each control strategy.

Table 4. Comparison between the Simulation Results for the Crystal Size in (110) Direction under Three Different Control Strategies for the Desired Growth Rate Ratio of $\langle \alpha \rangle = 1.11$

control strategy	$\langle h_{110} \rangle$	r_{10}	r_{50}	r_{90}	span
MPC	22.76	4.54	22.29	42.57	1.71
CTC: $T = 11\ °C$	25.14	4.45	24.58	46.82	1.72
CTC: $T = 13\ °C$	21.03	4.08	20.79	38.38	1.65
CSC: $\sigma = 3.34$	22.39	4.71	22.37	40.07	1.58
CSC: $\sigma = 3.52$	28.23	5.84	28.21	50.63	1.59

to a severe concentration drop. The relatively high span value for the MPC indicates a wide size distribution (high polydispersity) which is resulting from the significant drop in the concentration described previously.

In summary, MPC successfully drives the final crystal shape distribution to a desired set-point value and is found to be robust with respect to both an undesirable nucleation rate increased in the earlier stage and a drastic drop in the solute concentration. Additionally, a low polydispersity can be achieved depending on the desired crystal morphology. For instance, for $\langle \alpha \rangle = 1.11$, it is more likely to obtain a narrow crystal shape distribution with a high polydispersity because the crystallizer promptly responds to the depletion in the solute concentration which is reflected as a series of rise-and-fall in the crystallizer temperature in Figure 9, and the high polydispersity is resulting from uniformly nucleated crystals. For $\langle \alpha \rangle = 0.85$, however, crystals with desired shape and a low polydispersity can be achieved at a very high initial temperature in the crystallizer, because the system reaches its optimal temperature fast which results in the crystals uniformly nucleated along the batch.

■ CONCLUSIONS AND FUTURE WORK

The present work focuses on comparing the performance of the proposed model predictive control (MPC) scheme to that of constant temperature control (CTC) and constant super-saturation control (CSC) in regulating crystal shape and size distributions to desired values. More specifically, the CTC operating strategy drives the crystallizer conditions from the labile zone to the metastable zone, and the CSC policy maintains

the crystallizer conditions in the metastable zone, respectively, to maximize the size of protein crystals.

First, we focus on the modeling of a batch crystallization process used to produce tetragonal hen egg white lysozyme crystals via kinetic Monte Carlo (kMC) simulation. The kMC simulation models batch protein crystallization via adsorption, desorption, and migration mechanisms on the (110) and (101) faces. Then, in order to describe the nucleation occurring at different times in the batch simulations, the nucleation rate expression was extracted from experimental results by Galkin et al.[42] In addition, the dependence of the crystal growth on temperature and protein solute concentration is demonstrated in 3-D nonlinear models constructed from open-loop kMC simulations. The present work also develops mass and energy balances to account for the depletion in the protein solute concentration and the drop in the crystallizer temperature by crystallization. Finally, an MPC, which makes use of the mass and energy balances, is designed to produce crystals with a desired morphology by regulating the crystal growth conditions in the crystallizer through the manipulation of the jacket temperature which is in accordance with standard batch crystallization practice.

Simulation results show that the proposed MPC is able to regulate the crystal shape distribution to a desired set-point value while reducing the effect of an undesirable biased nucleation in the earlier stage and a drastic drop in the solute concentration. Comparing the simulation results of the MPC with those of other conventional operating strategies, crystals with a low polydispersity can be produced depending on the desired crystal morphology. For instance, for $\langle \alpha \rangle = 1.11$, the crystallizer under MPC results in a narrow crystal shape distribution with a high polydispersity, because the crystallizer immediately responds to the depletion in the solute concentration and produces a series of rise-and-fall in the jacket temperature computed from the proposed MPC in Figure 9, and a high polydispersity is resulting from uniformly nucleated crystals. For $\langle \alpha \rangle = 0.85$, however, crystals with desired morphology and a low polydispersity can be achieved at a very high initial temperature in the crystallizer, because the system reaches its optimal temperature fast and it causes a uniform nucleation rate along the batch. In this case, therefore, we can reduce the batch time considerably, because the system reaches its optimal state faster. Furthermore, only protein solute concentration and temperature measurements are needed to implement this operating policy in practice; no additional measurements such as crystal size and shape are required in the controller.

For the future work, crystal growth from a defect or the dislocation of a monomer in a crystal layer can be considered over the course of the kMC simulation run which may initiate a spiral crystal growth due to a fast migration process. At a low supersaturation range where it is observed in experiments that spiral growth dominates,[49] this consideration can potentially produce more realistic kMC simulation for the crystal growth.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: makis@seas.ucla.edu.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Rosenberger, F.; Vekilov, P. G.; Muschol, M.; Thomas, B. R. Nucleation and crystallization of globular proteins—What we know and what is missing. *J. Cryst. Growth* **1996**, *168*, 1−27.

(2) Vekilov, P. G.; Chernov, A. A. The physics of protein crystallization. *Solid State Phys.* **2003**, *57*, 1−147.

(3) Wiencek, J. M. New strategies for protein crystal growth. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 505−534.

(4) Galkin, O.; Vekilov, P. G. Direct determination of the nucleation rates of protein crystals. *J. Phys. Chem. B* **1999**, *103*, 10965−10971.

(5) Pusey, M. L.; Nadarajah, A. A model for tetragonal lysozyme crystal nucleation and growth. *Cryst. Growth Des.* **2002**, *2*, 475−483.

(6) Durbin, S. D.; Feher, G. Crystal growth studies of lysozyme as a model for protein crystallization. *J. Cryst. Growth* **1986**, *76*, 583−592.

(7) Forsythe, E. L.; Nadarajah, A.; Pusey, M. L. Growth of (101) faces of tetragonal lysozyme crystals: Measured growth-rate trends. *Acta Cryst. D* **1999**, *55*, 1005−1011.

(8) Kurihara, K.; Miyashita, S.; Sazaki, G.; Nakada, T.; Suzuki, Y.; Komatsu, H. Interferometric study on the crystal growth of tetragonal lysozyme crystal. *J. Cryst. Growth* **1996**, *166*, 904−908.

(9) Simon, L.; Oucherif, K.; Nagy, Z.; Hungerbuhler, K. Histogram matching, hypothesis testing, and statistical control-chart-assisted nucleation detection using bulk video imaging for optimal switching between nucleation and seed conditioning steps. *Ind. Eng. Chem. Res.* **2010**, *49*, 9932−9944.

(10) Simon, L.; Oucherif, K.; Nagy, Z.; Hungerbuhler, K. Bulk video imaging based multivariate image analysis, process control chart and acoustic signal assisted nucleation detection. *Chem. Eng. Sci.* **2010**, *65*, 4983−4995.

(11) Nayhouse, M.; Kwon, J. S.; Christofides, P. D.; Orkoulas, G. Crystal shape modeling and control in protein crystal growth. *Chem. Eng. Sci.* **2013**, *87*, 216−223.

(12) Kwon, J. S.; Nayhouse, M.; Christofides, P.; Orkoulas, G. Modeling and control of protein crystal shape and size in batch crystallization. *AIChE J.* **2013**, *59*, 2317−2327.

(13) Durbin, S. D.; Feher, G. Simulation of lysozyme crystal growth by the Monte Carlo method. *J. Cryst. Growth* **1991**, *110*, 41−51.

(14) Müller, C. How to describe protein crystals correctly? Case study of lysozyme crystals. Ph.D. Thesis, Martin-Luther-Universitat Halle-Wittenberg, 2012.

(15) Fichthorn, K. A.; Weinberg, W. H. Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.* **1991**, *95*, 1090−1096.

(16) Müller-Krumbhaar, H.; Binder, K. Dynamic properties of the Monte Carlo method in statistical mechanics. *J. Stat. Phys.* **1973**, *8*, 1−24.

(17) Bortz, A. B.; Kalos, M. H.; Lebowitz, J. L. New algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.* **1975**, *17*, 10−18.

(18) Dai, J.; Kanter, J. M.; Kapur, S. S.; Seider, W. D.; Sinno, T. On-lattice kinetic Monte Carlo simulations of point defect aggregation in entropically influenced crystalline systems. *Phys. Rev. B* **2005**, *72*, 134102.

(19) Dai, J.; Seider, W. D.; Sinno, T. Coarse-grained lattice kinetic Monte Carlo simulation of systems of strongly interacting particles. *J. Chem. Phys.* **2008**, *128*, 194705.

(20) Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **1976**, *22*, 403−434.

5013

dx.doi.org/10.1021/ie400584g | *Ind. Eng. Chem. Res.* 2014, 53, 5002−5014

(21) Gillespie, D. T. Exact stochastic simulation of chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340−2361.

(22) Gillespie, D. T. Monte Carlo simulation of random walks with residence time dependent transition probability rates. *J. Comput. Phys.* **1978**, *28*, 395−407.

(23) Gillespie, D. T. A rigorous derivation of the chemical master equation. *Phys. A* **1992**, *188*, 404−425.

(24) Gillespie, D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **2001**, *115*, 1716−1733.

(25) Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **2007**, *58*, 35−55.

(26) Rathinam, M.; Petzold, L. R.; Cao, Y.; Gillespie, D. T. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chem. Phys.* **2003**, *119*, 12784−12794.

(27) Reese, J. S.; Raimondeau, S.; Vlachos, D. G. Monte Carlo algorithms for complex surface reaction mechanisms: efficiency and accuracy. *J. Comput. Phys.* **2001**, *173*, 302−321.

(28) Snyder, M. A.; Chatterjee, A.; Vlachos, D. G. Net-event kinetic Monte Carlo for overcoming stiffness in spatially homogeneous and distributed systems. *Comput. Chem. Eng.* **2005**, *29*, 701−712.

(29) Christofides, P. D.; Armaou, A.; Lou, Y.; Varshney, A. *Control and Optimization of Multiscale Process Systems*; Birkhäuser: Boston, 2008.

(30) Chiu, T.; Christofides, P. D. Robust control of particulate processes using uncertain population balances. *AIChE J.* **2000**, *46*, 266−280.

(31) El-Farra, N.; Chiu, T.; Christofides, P. D. Analysis and control of particulate processes with input constraints. *AIChE J.* **2001**, *47*, 1849−1865.

(32) Ke, S. C.; DeLucas, L. J.; Harrison, J. G. Computer simulation of protein crystal growth using aggregates as the growth unit. *J. Phys. D: Appl. Phys.* **1998**, *31*, 1064−1070.

(33) Cacioppo, E.; Munson, S.; Pusey, M. L. Protein solubilities determined by a rapid technique and modification of that technique to a micro-method. *J. Gryst. Growth* **1991**, *110*, 66−71.

(34) Cacioppo, E.; Pusey, M. L. The solubility of the tetragonal form of hen egg white lysozyme from pH 4.0 to 5.4. *J. Gryst. Growth* **1991**, *114*, 286−292.

(35) Feher, G.; Kam, Z. Nucleation and growth of protein crystals: General principles and assays. *Methods Enzym.* **1985**, *114*, 77−1112.

(36) Gilmer, G. H.; Bennema, P. Simulation of crystal growth with surface diffusion. *J. Appl. Phys.* **1972**, *43*, 1347.

(37) Shi, D.; Mhaskar, P.; El-Farra, N. H.; Christofides, P. D. Predictive control of crystal size distribution in protein crystallization. *Nano-technology* **2005**, *16*, S562−S574.

(38) Christofides, P. D. *Model-Based Control of Particulate Processes*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2002.

(39) Nanev, C. N.; Tsekova, D. Heterogeneous nucleation of hen-egg-white lysozyme-molecular approach. *Cryst. Res. Technol.* **2000**, *35*, 189−195.

(40) Suzuki, Y.; Miyashita, S.; Komatsu, H.; Sato, K.; Yagi, T. Crystal growth of hen egg white lysozyme under high pressure. *Jpn. J. Appl. Phys.* **1994**, *33*, 1568−1570.

(41) Miller, S. Modeling and quality control strategies for batch cooling crystallizers. Ph.D. Thesis, The University of Texas at Austin, Texas, 1993

(42) Galkin, O.; Vekilov, P. G. Nulceation of protein crystals: Critical nuclei, phase behavior, and control pathways. *J. Cryst. Growth* **2001**, *232*, 63−76.

(43) Aldabaibeh, N.; Jones, M. J.; Myerson, A.; Ulrich, J. The solubility of orthorhombic lysozyme chloride crystals obtained at high pH. *Cryst. Growth Des.* **2009**, *9*, 3313−3317.

(44) Aldabaibeh, N. Towards an improved understanding of the variables controlling crystal modification and phase transition in Lysozyme crystals. Dissertation, Illinois Institute of Technology, 2010.

(45) Weber, M.; Jones, M. J.; Ulrich, J. Crystallization as a purification method for jack bean urease: On the suitability of poly(ethylene), Li_2SO_4 and NaCl as precipitants. *Cryst. Growth Des.* **2008**, *8*, 711−716.

(46) Müller, C.; Liu, Y.; Migge, A.; Pietzsch, M.; Ulrich, J. Recombinant L-Asparaginase B and its crystallization—What is the nature of protein crystals? *Chem. Eng. Technol.* **2011**, *34*, 571−577.

(47) Müller, C.; Ulrich, J. A more clear insight of the lysozyme crystal composition. *Cryst. Res. Technol.* **2011**, *46*, 646−650.

(48) Shi, D.; El-Farra, N. H.; Li, M.; Mhaskar, P.; Christofides, P. D. Predictive control of particle size distribution in particulate processes. *Chem. Eng. Sci.* **2006**, *61*, 268−281.

(49) Zhang, Y.; Doherty, M. F. Simultaneous prediction of crystal shape and size for solution crystallization. *AIChE J.* **2004**, *50*, 2101−2112.

5014

dx.doi.org/10.1021/ie400584g | *Ind. Eng. Chem. Res.* 2014, 53, 5002−5014