

# Characterization of Propionyl Phosphate Hydrolysis Kinetics by Data-Rich Experiments and In-Line Process Analytical Technology

Cuixian Yang,\* Hanzhou Feng, and Kevin Stone



Cite This: *Org. Process Res. Dev.* 2021, 25, 507–515



Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

**ABSTRACT:** In-depth characterization of reaction kinetics often requires a considerable amount of experimental results under various conditions. Recent advances in data-rich experimentation enable the collection of sufficient data to investigate reactions with only a limited number of experiments. In this study, we developed a cost-efficient, robust approach by utilizing data-rich experimentation to characterize propionyl phosphate hydrolysis reaction kinetics. Specifically, an Fourier transform infrared (FTIR)-based process analytical technology (PAT) and off-line NMR calibration allowed the establishment of a quantitative FTIR multivariant model. This PAT was then integrated with repeated temperature scanning (RTS) to generate a massive database in a single experiment. The data were subsequently used for kinetic analysis, and two key characteristic reaction parameters (the activation energy and pre-exponential factor) were determined on the basis of the assumption of first-order kinetics. We envision that the integrative platform developed in this study can be broadly applied to investigations of the kinetics of a wide range of similar liquid-phase reactions.

**KEYWORDS:** *propionyl phosphate, hydrolysis, FTIR, reaction kinetics, process analytical technology*

## 1. INTRODUCTION

Recent years have witnessed the rapid development and application of enzyme-catalyzed reactions in the pharmaceutical industry, as they offer benefits of unparalleled selectivity, increased atom economy, tunability, and improved safety.<sup>1</sup> These enzyme-based reactions have been widely used in syntheses of high-value products, such as nucleoside-based drugs for the treatment of different viral infections.<sup>2,3</sup>

Adenosine-5'-triphosphate (ATP)-dependent enzymatic reactions are ubiquitous in nature and have been extensively employed for organic synthesis. In such enzymatic reactions, ATP needs to be used in catalytic amounts and regenerated in situ to minimize expense and simplify isolation of the products. ATP regeneration from adenosine-5'-diphosphate (ADP), adenosine-5'-monophosphate (AMP), or adenosine by using inexpensive phosphate donors has been explored in the past<sup>4–7</sup> and provides high longevity, compatibility, and robustness under process conditions. Common phosphate donors include acetyl phosphate (AcP), phosphoenolpyruvate (PEP), and methoxycarbonyl phosphate [ $\text{CH}_3\text{OC}(\text{O})\text{OPO}_3^{2-}$ , MCP].<sup>8,9</sup> For practical-scale organic synthesis, there are currently three different procedures for enzymatic ATP regeneration.<sup>8,9</sup> The first procedure uses AcP as the phosphorylating agent and acetate kinase as the catalyst; the second uses PEP and pyruvate kinase; and the third uses MCP and acetate kinase.

To better understand the enzymatic phosphorylation reaction, it is necessary to know the reaction mechanism and chemical stability of the phosphate donor as one key raw material for in situ ATP regeneration. AcP is one of most commonly used phosphate donors for a wide range of enzymatic reactions, and its hydrolysis mechanism and associated reaction kinetics have been widely studied.<sup>10–12</sup>

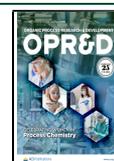
AcP degradation was reported to follow first-order kinetics in many cases studied by the Koshland group,<sup>10</sup> including those involving various pH conditions and the presence of a metal ion (magnesium) and a nucleophilic reagent (pyridine). It was also found that the hydrolysis rate is relatively unaffected by changes in pH in the near-neutrality region (pH 5–9), although it is greatly accelerated in strongly acidic or strongly basic environments.<sup>10</sup> Also, there have been studies investigating the specific location of the oxygen bond broken during the hydrolysis reaction of AcP. For example, the acid- and base-catalyzed hydrolysis of AcP reactions have been shown to take place by cleavage of the C–O bonds, while the uncatalyzed reaction and the pyridine-catalyzed reaction proceed by O–P bond splitting.<sup>13</sup>

In one drug development project at Merck & Co., Inc. (Kenilworth, NJ, USA), enzymatic phosphorylation is used as an important step for the synthesis of an active pharmaceutical ingredient. Propionyl phosphate (PrP) was chosen as the phosphate donor because of the benefits for the downstream biocatalytic process. However, the degradation kinetics of PrP has not been extensively reported before. Without appropriate process control, PrP hydrolysis in aqueous solution can be a non-negligible reaction that competes with the desired enzymatic reaction. Understanding of the PrP degradation kinetics is of great importance for risk mitigation during scale-

**Special Issue:** Celebrating Women in Process Chemistry

**Received:** October 5, 2020

**Published:** February 17, 2021



up process development. In particular, the time duration in a large-scale process can easily be an order of magnitude higher in than the lab process. Any unplanned event can significantly extend the process of charging key reagents/enzymes to the reaction vessel and delay triggering of the desired enzymatic reaction. This can subsequently result in degradation of the phosphate donor PrP, which eventually leads to incomplete reaction, increased impurities, and poor yields.

The temperature-dependent hydrolysis reaction cannot be stopped easily, therefore making it challenging to use traditional off-line analytical tools, such as HPLC, to capture the accurate real-time concentration profiles. In the past few years, the concept of process analytical technology (PAT) has been promoted by the U.S. Food and Drug Administration to design, analyze, and control manufacturing through timely measurements of critical quality and performance attributes of intermediate and output materials.<sup>14</sup> The goal of PAT is to enhance understanding and control of the pharmaceutical manufacturing process. The PAT framework primarily comprises efficient process analyzers and control tools as well as multivariate tools for design, data acquisition, and data analysis. The gained information can be used to develop first-principles models and elevate process understanding. Fourier transform infrared (FTIR) spectroscopy is a robust analytical method popularly used as PAT. FTIR spectra are determined by the molecular structure on the basis of the energies of chemical bond vibrations,<sup>15</sup> and recording a spectrum typically takes only a few seconds, making this method well-suited for monitoring real-time chemical changes in the reaction mixture. Consequently, it provides an exceptional analytical option to efficiently characterize reaction behavior. Notably, FTIR spectroscopy has already been demonstrated for the quantitative monitoring of hydrolysis reactions *in situ*.<sup>16,17</sup>

To study reaction kinetics, people traditionally rely on multiple isothermal reaction experiments at a number of different temperatures to develop an Arrhenius relationship for the observed rate constants in liquid or multiphase reactions. The activation energy ( $E_a$ ) of the reaction is then calculated from the slope of the “fitted” straight line through a limited number of data points. In comparison, isoconversional methods for the kinetic analysis of non-isothermal data, mainly including the Ozawa–Flynn–Wall,<sup>18</sup> Kissinger–Akahira–Sunose,<sup>19</sup> and Friedman–Ozawa<sup>20</sup> methods, have been applied to give meaningful and highly reliable kinetic parameters in a wide range of circumstances. However, multiple experimental runs are still needed for the isoconversional methods, which could be time- and resource-consuming and sometimes require extra expensive materials. In 2000, Ozawa<sup>21</sup> proposed and demonstrated the feasibility of the “repeated temperature scanning” (RTS) method (also called the “quasi-isoconversional” method) to obtain kinetic parameters from a single time-saving and data-rich experiment. The mode of temperature change is neither necessarily uniform nor controlled precisely, so there is sufficient flexibility in experimental execution. This method was further introduced in the form of a temperature scanning reactor by Wojciechowski,<sup>22</sup> primarily for studying gas-phase flow reactions in the petrochemical industry. To date, assisted by various powerful *in situ* reaction monitoring tools, this method has been used in batch<sup>23,24</sup> and flow reaction studies.<sup>25,26</sup>

In this study, we adopted the RTS experimental method to study the hydrolysis of PrP. Combined with *in situ* FTIR monitoring of the reaction extent and DynoChem modeling

software, RTS was utilized for kinetic analysis of the PrP hydrolysis reaction in a rapid, efficient, and robust manner. The findings of this work show that this approach was highly effective and provided considerable amounts of data using a single well-designed experimental run. Thus, this work demonstrates the potential of our approach for wide application to similar reactions.

## 2. EXPERIMENTAL SECTION

**2.1. Materials and Methods.** Bis-Tris buffer (50 mM, pH 8) was prepared by dissolving Bis-Tris free base (Sigma-Aldrich, lot no. SLBZ8488) in water (Thermo Scientific, lot no. 191946), titrated with propionic acid (ThermoFisher Scientific, lot no. R07F026) to pH  $\sim$ 8, and kept at 4 °C until being used. The Bis-Tris buffer was used as the solvent for the hydrolysis reaction of propionyl phosphate monoammonium (PrP mono) to mimic the enzymatic reaction conditions.

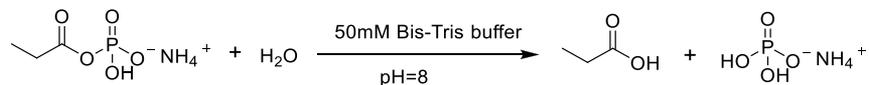
An EasyMax 102 advanced thermostat system was used to carry out the reaction. To a 100 mL EasyMax vessel (preset jacket temperature  $T_j = 5$  °C) was charged 20 mL of the prepared 50 mM Bis-Tris (pH 8) followed by 3.0454 g of solid PrP mono (Shanghai SynTheAll, lot no. CR-C19041282-B219001, 91.3% purity), and the solution was flushed with another 10 mL of 50 mM Bis-Tris (pH 8). The pH of the resulting PrP mono solution was around 3.6 because of the acidity of PrP mono salt. The solution pH was adjusted to  $\sim$ 8.2 by addition of  $\sim$ 2.6 mL of 37% KOH (Fisher, lot no. 166546 SP226-1). The stirring speed was set at 500 rpm. The initial PrP mono concentration was around 0.49 mol/L. The RTS experiment was designed and applied to this hydrolysis reaction. The reactor internal temperature ( $T_r$ ) was programmed to increase from 5 to 40 °C over 5 h and then decrease from 40 to 5 °C over 5 h. The same cycle was repeated four times, including a few temperature holding stages. The detailed temperature profiles are shown in Figure 3a.

The pH meter and thermocouple were inserted into the reactor in order to monitor the reaction parameters. The reaction extent was monitored by FTIR spectroscopy (Mettler-Toledo Fiber MultiplexIR) in real time. The IR probe (Mettler-Toledo DiComp) was immersed in the reaction medium, and the spectra were collected at 1 min intervals over the entire reaction course. The spectral range was 2500–650  $\text{cm}^{-1}$  and the spectral resolution was 8  $\text{cm}^{-1}$ . iControl software was paired with the EasyMax system to control the experiments and also enabled scientists to gain an in-depth understanding of the reaction data.

Seven off-line samples were taken during the reaction and subjected to quantitative  $^{31}\text{P}$  NMR analysis on a Bruker 500 MHz NMR spectrometer to determine the concentrations of the starting material (PrP mono) and the product (free phosphate). For NMR tests, approximately 300  $\mu\text{L}$  of reaction solution was added into  $\sim$ 1 mL of  $\text{D}_2\text{O}$  (CAS no. 7789-20-0, Acros Organics, lot no. B0758205, 99.8 atom %) with  $\sim$ 40 mg of tetraphenylphosphonium chloride ( $(\text{C}_6\text{H}_5)_4\text{PCl}$ , CAS no. 2001-45-8, Sigma-Aldrich, 98 wt %) as an internal P standard. The NMR spectra were analyzed using MestReNova Chemists software to determine the concentrations of off-line samples, which were integrated with the on-line IR data to establish the multivariate model. (A representative  $^{31}\text{P}$  NMR spectrum is shown in Figure S1).

The concentration profiles obtained by FTIR spectroscopy were fit to a first-order reaction kinetics model and the

## Scheme 1. Balanced Chemical Equation for the Hydrolysis of PrP Mono



Arrhenius equation using DynoChem modeling software to obtain  $E_a$  and the reference rate constant ( $k_{\text{ref}}$ ) at the reference temperature ( $T_{\text{ref}}$ ) of 33 °C (the pre-exponential factor  $A$  can be obtained from  $k_{\text{ref}}$ ). The theory behind the kinetic analysis is presented in the following section.

**2.2. Theory.** 2.2.1. *The Hydrolysis Reaction.* The overall hydrolysis of PrP mono can be represented as shown in Scheme 1.

Typically, the hydrolysis reaction mechanism includes three steps: addition, elimination, and proton transfer.<sup>11,24</sup> As the addition step is usually the rate-controlling step, the overall rate can be written as

$$-\frac{dC_A}{dt} = k' C_A C_B \quad (1)$$

where  $C_A$  is the concentration of PrP mono and  $C_B$  is the concentration of water. Under conditions with an excess of water, a pseudo-first-order reaction can be assumed, and eq 1 simplifies to

$$-\frac{dC_A}{dt} = k C_A \quad (2)$$

The rate constant  $k$  is expressed by the Arrhenius equation as

$$k = A \exp\left(-\frac{E_a}{RT}\right) \quad (3)$$

in which  $E_a$  is the activation energy,  $A$  is the pre-exponential factor,  $R$  is the universal gas constant, and  $T$  is the absolute temperature.

2.2.2. *Kinetic Parameter Evaluation.* The reaction kinetic parameters of this work were obtained using the RTS experiment and DynoChem model analysis. Specifically, the reaction temperature was controlled to go through the designed heating and cooling steps in repeated cycles, as shown in Figure 3a. The experimental data, including the time ( $t$ ), temperature ( $T$ ), and conversion ( $\alpha$ ), were used for analysis. Importantly, the Arrhenius equation was considered to be valid for the entire temperature change process.

Also, it was assumed that the rate of the process,  $d\alpha/dt$ , is dependent on  $T$  and  $\alpha$  and can be expressed as a product of two independent functions, an  $\alpha$ -dependent function  $A'(\alpha)$  and a  $T$ -dependent function  $g(T)$ , as shown in eq 4:

$$\frac{d\alpha}{dt} = F(T, \alpha) = A(\alpha) \exp\left(-\frac{E_a}{RT}\right) f(\alpha) = A'(\alpha) g(T) \quad (4)$$

where  $A'(\alpha) = A(\alpha)f(\alpha)$  and  $g(T) = \exp(-E_a/RT)$ . The function  $f(\alpha)$  reflects the mechanism/reaction model of the process.

Unlike the previous studies using the Ozawa–Friedman plot to obtain  $E_a$  and  $A$ , this work used DynoChem to calculate these kinetic parameters. Specifically, the collected  $\alpha$ ,  $t$ , and  $T$  experimental data were fed into the model described by eq 4 to solve for  $E_a$  and  $A$  using a numerical method as described in section 2.3.2.

**2.3. Modeling.** 2.3.1. *Quantitative IR Models.* FTIR models of the analyte of interest were developed by univariate

and multivariate modeling approaches. A univariate model can be calculated via linear regression by relating the absorbance values at a single wavenumber to the concentrations determined by <sup>31</sup>P NMR spectroscopy.

Principal component analysis (PCA) and interval partial-least-squares regression (iPLS) were used to build multivariate qualitative and quantitative models. PCA is an unsupervised variable reduction technique used to construct new variables, known as principal components (PCs). The principle of iPLS is to split the spectra into a given number of equidistant subintervals and develop a PLS regression model for each subinterval.<sup>27,28</sup> Cross-validation was performed for each of these models, and the subinterval that provided the lowest cross-validation modeling error was selected. The model computation was achieved using MATLAB (R2015a) equipped with PLS\_Toolbox (version 8.7). The multivariate model performance was assessed using the root-mean-square error of cross-validation (RMSECV) and root-mean-square error of calibration (RMSEC), which is calculated as

$$\text{RMSEC} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_{i,\text{pred}} - y_{i,\text{ref}})^2}{n - m - 1}} \quad (5)$$

where  $n$  is the sample size,  $m$  is the number of latent variables included in the iPLS model,  $\hat{y}_{i,\text{pred}}$  is the concentration of the  $i$ th sample predicted by the iPLS model, and  $y_{i,\text{ref}}$  is the concentration of the  $i$ th sample determined by <sup>31</sup>P NMR spectroscopy. Both univariate and multivariate models were subsequently used to provide real-time concentration profiles of PrP mono and phosphate.

2.3.2. *DynoChem Model.* The limited number of concentration data points from NMR analysis of the seven off-line samples is not sufficient to establish a reliable DynoChem model. Instead, the limited NMR data were used to calibrate an in situ FTIR dataset of up to ~3000 data points. To determine the optimum sampling from this large data set for the estimation of the kinetic parameters, different numbers of data points were tested in the regression of the DynoChem model.

The thermodynamic parameters of the reaction in section 2.2.1, namely, the reference pseudo-first-order rate constant ( $k_{\text{ref}}$ ) and the associated activation energy ( $E_a$ ), were regressed using DynoChem 5 process modeling software. This regression is a maximum-likelihood estimation that can be represented as the maximization of the log likelihood as expressed in eq 6:

$$-\ln(L) = 2N \ln(2\pi) + \frac{1}{2} \sum_{j=1}^{n_j} \sum_{i=1}^{n_i} \ln(\sigma_{ij}^2) + \frac{1}{2} \sum_{j=1}^{n_j} \sum_{i=1}^{n_i} \frac{[y_{ij}^{\text{exp}} - y_{ij}(\theta_k)]^2}{\sigma_{ij}^2} \quad (6)$$

where  $L$  is the likelihood of an observation,  $N$  is the number of observations,  $\sigma_{ij}$  is the standard deviation of the normal distribution of the measurement error,  $y_{ij}^{\text{exp}}$  is the experimental measurement, and  $y_{ij}(\theta_k)$  is the corresponding model prediction depending on  $k$  model parameters  $\theta$ .

The maximization of the log likelihood can be simplified by minimization of the third term in eq 6, simplified as the  $\chi^2$  expression as shown in eq 7:

$$\chi^2 = \sum_{j=1}^{n_j} \sum_{i=1}^{n_i} \frac{[y_{ij}^{\text{exp}} - y_{ij}(\theta_k)]^2}{\sigma_{ij}^2} \quad (7)$$

The DynoChem solver minimizes the  $\chi^2$  value, or SSQ, by manipulating  $y_{ij}^{\text{exp}}$  through the parameters of the model  $\theta_k = [k_{\text{ref}}, E_a]$  with the gradient-based Levenberg–Marquardt routine.<sup>29,30</sup> The individual iterations are evaluations of the model performed with the DynoChem simulator, which uses a Rosenbrock integration method to solve the mass balance differential and algebraic equations associated with the reaction in section 2.2.1. The model values  $y_{ij}(\theta_k)$  that are selected from these equations for simulation and fitting are the concentrations of the substrate (PrP mono) and the product (free phosphate).

For all regressions in this work, the variance of the measurement error is estimated by a relative weighting method as expressed in eq 8:

$$\sigma_{ij}^2 = \left( \frac{n_j}{n_T} \right)^2 y_{j,\text{max}}^2 \quad (8)$$

where  $n_j$  is the number of points in component series  $j$ ,  $n_T$  is the total number of points in the experiment, and  $y_{j,\text{max}}^2$  is the square of the maximum measured value of component  $j$ . The first factor in this expression normalizes the weight of each component for regressions where there are multiple components with potentially distinct numbers of data points. We consider this model suitable to estimate the measurement error because we reasonably expect the absolute measurement error to increase with the magnitude of the measurement.

As described in section 2.2.1, the rate law for the reaction of interest can be simplified as shown in eq 2, where  $C_A$  is the concentration of PrP mono. The rate constant  $k$  in eq 2 at any given temperature is evaluated in DynoChem using the following modified Arrhenius expression:

$$\ln(k(T_j)) = \ln(k_{\text{ref}}(T_{\text{ref}})) - \frac{E_a}{R} \left( \frac{1}{T_{\text{ref}}} - \frac{1}{T_j} \right) \quad (9)$$

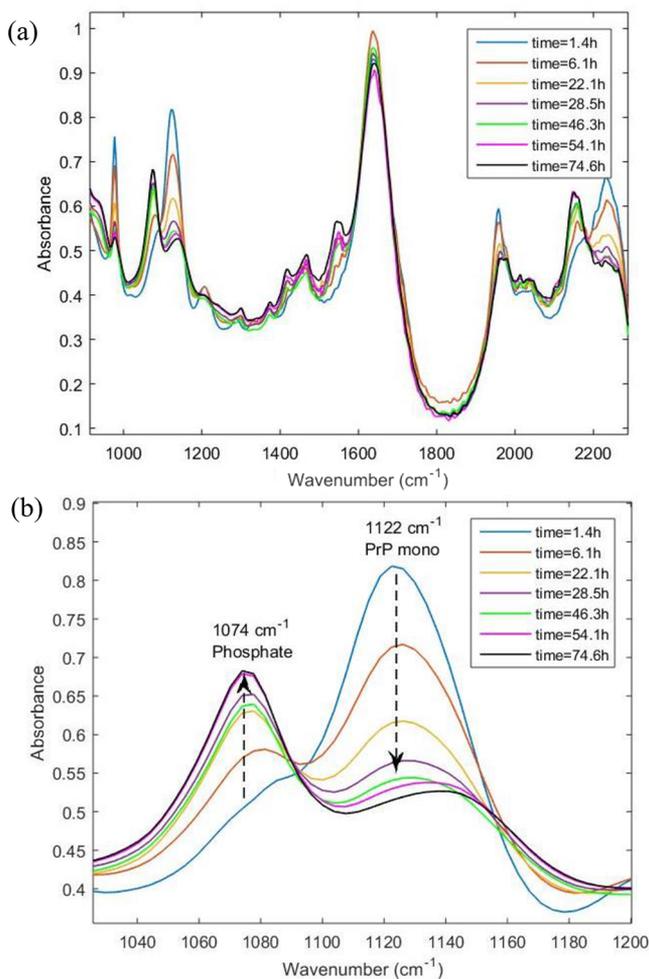
in which  $k(T_j)$  is the apparent rate constant of the reaction at temperature  $T_j$  and  $k_{\text{ref}}$  is the reference rate constant at the reference temperature  $T_{\text{ref}}$ . The two independent parameters used for fitting in the DynoChem regression are  $k_{\text{ref}}$  (where the choice of  $T_{\text{ref}}$  may vary depending on the reaction model and experiment) and  $E_a$ . Equation 9 is optimal for regression of the kinetics because collinearity between the apparent rate constant and activation energy is often minimized with an appropriate choice of reference temperature.

In a traditional regression of kinetic data for reactions collected under isothermal conditions, integration of the rate law in eq 2 for each experiment can be conducted with a single evaluation of  $k(T)$ . However, in this work we present a methodology by which non-isothermal data are used to regress the kinetic parameters of this reaction. In this case, DynoChem dynamically evaluates the rate constant  $k(T)$  according to eq 9 on the basis of the temperature data collected in the experiments and linear interpolations of those data where necessary. All of the temperature profiles used in these

experiments were linear in nature, so those interpolations did not contribute to error.

### 3. RESULTS AND DISCUSSION

**3.1. IR Spectra and Model Development.** Significant spectral changes can be found in the raw FTIR spectra shown in Figure 1a. The 1000–1200  $\text{cm}^{-1}$  spectral region presents



**Figure 1.** (a) Selected representative raw IR spectra of the reaction. (b) IR spectra of off-line NMR samples in the 1000–1200  $\text{cm}^{-1}$  region.

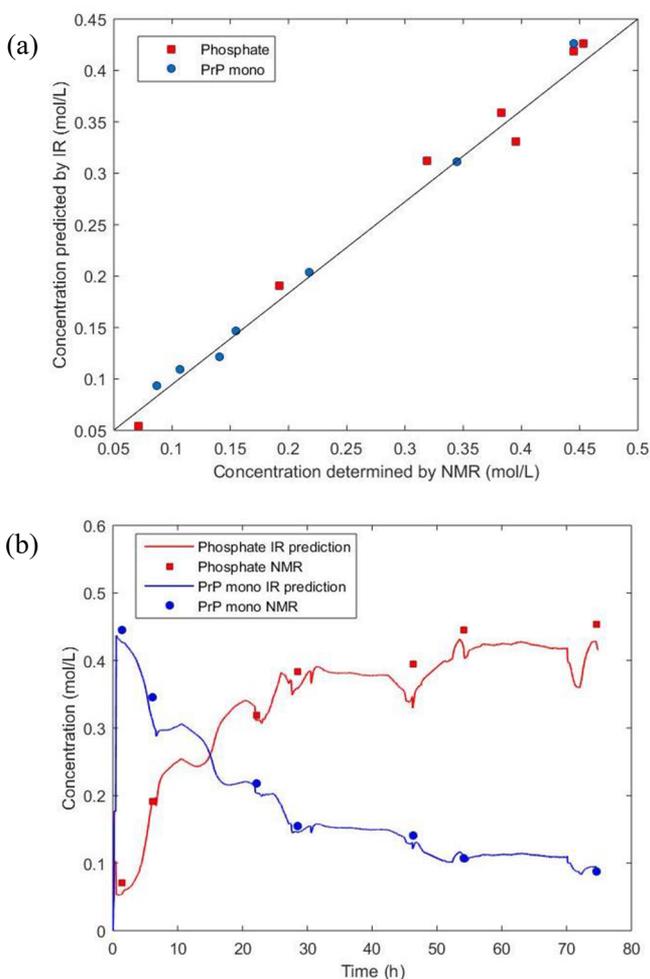
more evident changes associated with variation of the concentration. As can be seen in Figure 1b, the absorbance at 1074  $\text{cm}^{-1}$  increases as the reaction progresses, indicating the formation of the product phosphate, while there is a decrease in the absorbance at 1122  $\text{cm}^{-1}$ , which is associated with the disappearance of the starting material PrP mono. The identified regions were subsequently used to develop univariate linear regression models, as described by eqs 10 and 11:

$$y_1 = 1.097x_1 - 0.47 \quad (10)$$

$$y_2 = 2.075x_2 - 0.99 \quad (11)$$

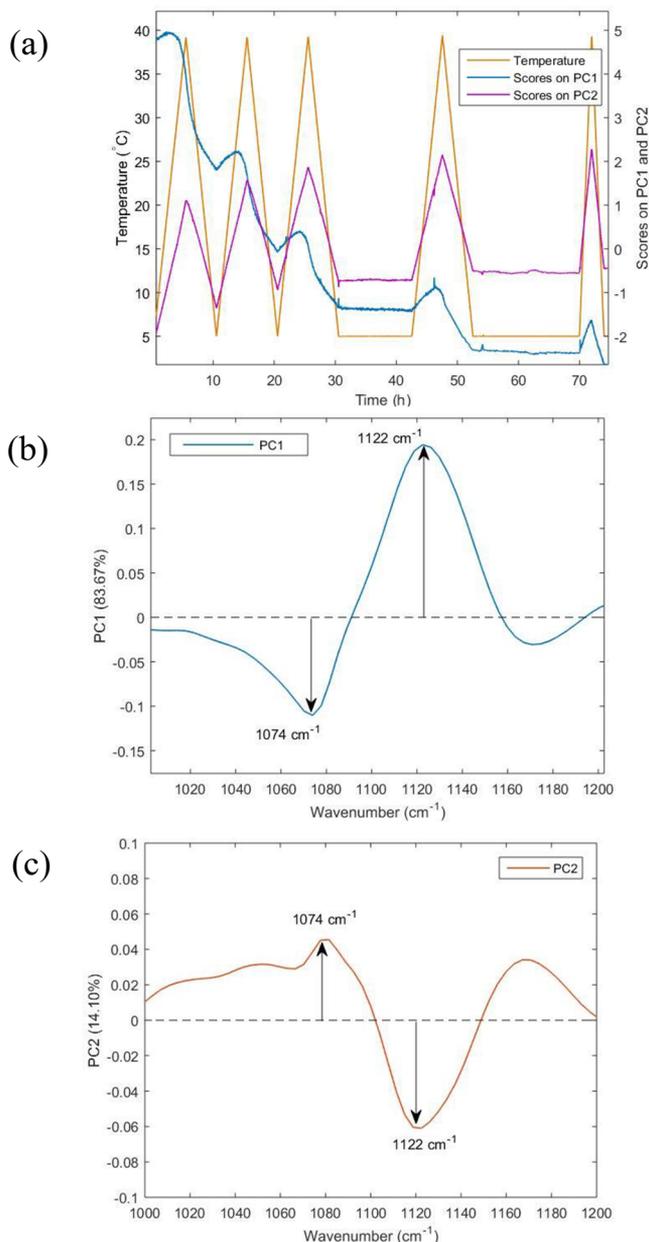
where  $x_1$  and  $x_2$  are the FTIR absorbances at 1122 and 1074  $\text{cm}^{-1}$ , respectively, and  $y_1$  and  $y_2$  are the predicted concentrations of PrP mono and phosphate, respectively. The univariate models of PrP mono and phosphate give RMSEC values of 0.022 and 0.036 mol/L, respectively.

Although the result of the model fitting shown in Figure 2a suggests reasonable linearity, Figure 2b shows that the prediction errors are quite significant and thus that the model performance is not optimal.



**Figure 2.** (a) Fitting results for the phosphate and PrP mono univariate models. (b) Real-time concentrations predicted by the univariate models vs off-line measured concentrations.

PCA was performed to probe the cause of the limited model performance as well as to explore opportunities for model improvement. The FTIR spectra were preprocessed by standard normal variate (SNV) and mean centering. The first and second principal components (PC1 and PC2) explain 83.7% and 14.1% of the spectra, respectively. In Figure 3a, the scores on PC1 reflect the concentration change during the reaction. On the other hand, the scores on PC2 exhibit a trend nearly identical to the temperature profile. The loading plots of PC1 and PC2 are displayed in Figure 3b and Figure 3c, respectively. The higher weight values of PC1 at 1074 and 1122  $\text{cm}^{-1}$  suggest that these regions have greater contributions to the overall concentration change, which justifies their use in the development of univariate models. The scores on PC2 are indicative of the temperature change. The higher weight values between 1060 and 1170  $\text{cm}^{-1}$  imply that this spectrum range is highly susceptible to temperature. The PCA results signify that the univariate models were not specific to the quantification of the concentrations because the wavenumbers used are substantially temperature-dependent (ex-

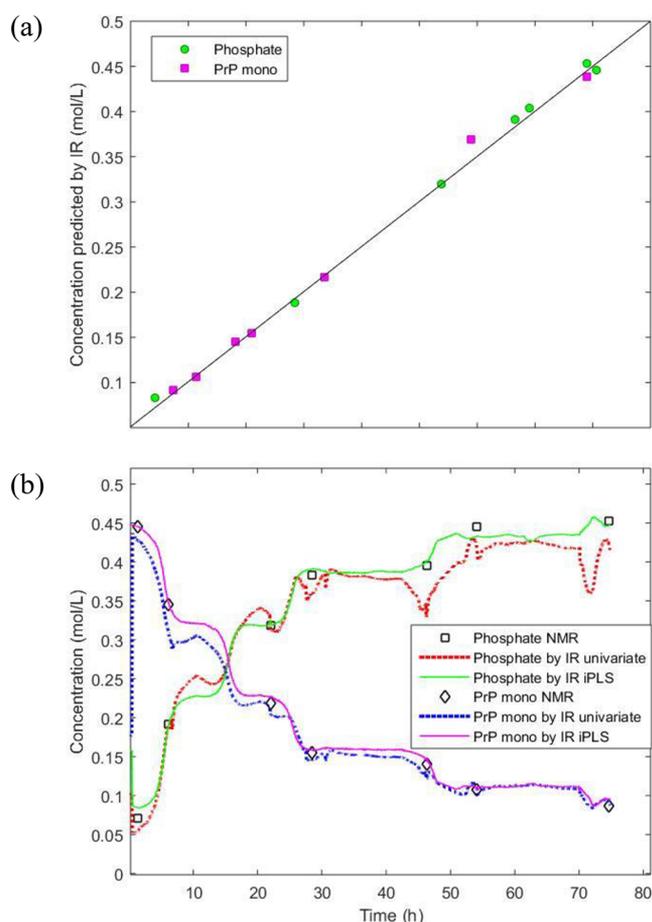


**Figure 3.** (a) Reaction temperature profile and PCA scores. (b) Loading plot of PC1. (c) Loading plot of PC2.

plained by PC2). Therefore, the univariate modeling approach is not robust and cannot provide adequate prediction performance for a system with considerable temperature change. Nevertheless, it should be noted that differential changes in the refractive indices of the FTIR diamond sensor and reaction materials with temperature may cause temperature-dependent variation. To this end, this study used a PLS algorithm that maximizes the covariance between the compound concentration and FTIR signal changes and is therefore less susceptible to the temperature difference. On the other hand, alternative analytical methods that are less susceptible to the temperature variation, such as UV transmittance, may provide a robust univariate model.

To improve the model performance and accuracy, multivariate models were developed by forward iPLS. The spectral range of 1010–1454  $\text{cm}^{-1}$  was included for building PrP mono and phosphate models. One latent variable was sufficient for

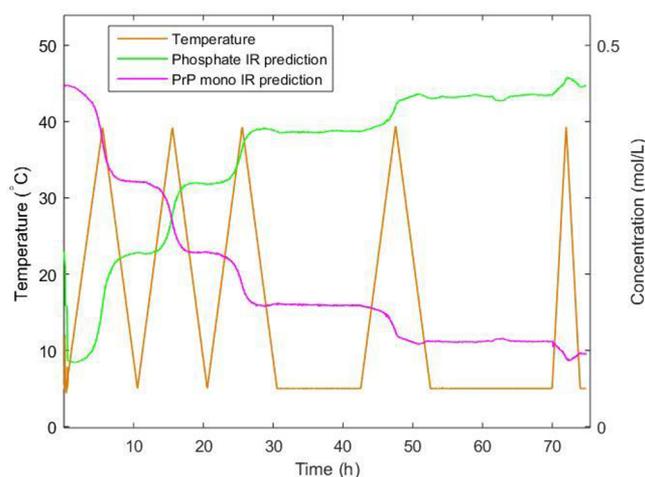
both models. The PrP mono model has an RMSEC of 0.0058 mol/L and an RMSECV of 0.010 mol/L, whereas the RMSEC and RMSECV of the phosphate model are 0.0085 and 0.017 mol/L, respectively. Multivariate modeling significantly improved the model performance and accuracy, as indicated by the better fitting results shown in Figure 4a and the



**Figure 4.** (a) Multivariate model fitting results for phosphate and PrP mono. (b) Comparison of IR predictions by the univariate and iPLS models.

considerably reduced prediction errors shown in Figure 4b. The real-time predictions by the iPLS model, displayed in Figure 5, are robust against the temperature variation.

Next, the real-time PrP mono and free phosphate concentrations predicted by the iPLS model along with temperature profile are displayed in Figure 5. The PrP mono concentration decreases and the free phosphate concentration increases as the temperature  $T_r$  increases from 5 to 38 °C in the first half of the temperature scanning cycle. The first derivative of the concentration,  $dC/dt$ , represents the reaction rate. The slopes of both curves (i.e., the reaction rate) increase with increasing  $T_r$ . At the maximum  $T_r$  of 38 °C,  $dC/dt$  appears to be the highest, which is rationalized by the Arrhenius equation. As  $T_r$  decreases from 38 to 5 °C in the second half of the temperature scanning cycle, reaction rate decreases, and the concentration shows no significant change. During two temperature holding periods (relative reaction time: 30–42 h and 52–70 h) at 5 °C, the concentrations of the starting material and product remain nearly unchanged, likely because of the fairly low reaction rate at such a low

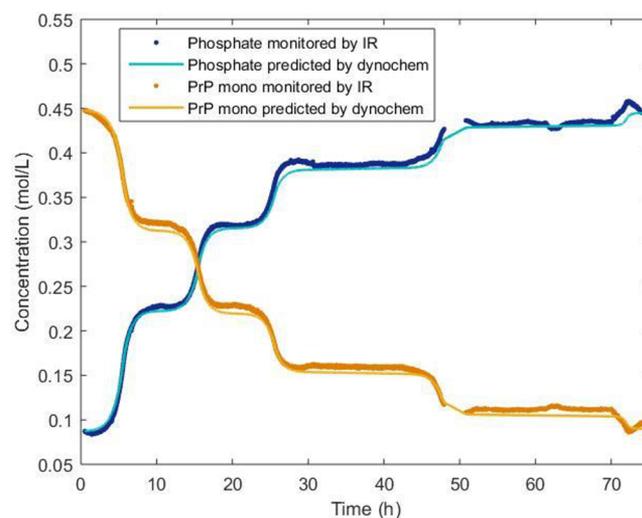


**Figure 5.** Time profiles of the real-time FTIR predictions of the concentrations by iPLS models and the reaction temperature change.

temperature. Five repeated temperature scanning cycles were applied to the reaction until the starting material (PrP mono) was mostly depleted and converted to product free phosphate. The measured pH value during the course of the dynamic experiment (Figure S10) was between 6 and 8.5, which is within the pH-independent region for such hydrolysis kinetics.<sup>10</sup>

**3.2. Kinetic Model Fitting.** **3.2.1. Obtaining the Activation Energy and Reference Rate Constant.** In this section, the full 3102 data sets of PrP mono/free phosphate concentration from IR monitoring (section 3.1) and the temperature (from iControl) along with the corresponding relative reaction times were applied to the kinetic model fitting to obtain two important kinetics parameters, i.e.,  $E_a$  and  $k_{ref}$  at  $T_{ref}$ . The kinetic model was established on the basis of the pseudo-first-order PrP mono hydrolysis kinetics assumption and the theory of the RTS method (sections 2.2 and 2.3.2).

As shown in Figure 6, the dotted lines are the PrP mono (orange) and free phosphate (blue) concentrations obtained from FTIR spectroscopy calibrated by off-line NMR analysis



**Figure 6.** Concentration profiles of phosphate and PrP mono estimated by IR (dotted lines) vs predictions of the DynoChem model (solid lines).

(section 3.1), while the solid curves represent the predicted concentration profiles after fitting to the kinetic model. The empty gap near 50 h in the IR profiles was due to data collection fouling. The confidence level was selected as 95% for fitting in DynoChem. The confidence intervals of both parameters obtained through pseudo-first-order kinetics model fitting are <10%, which is indicative of a suitable fit. Under the assumption that the fitting errors are normally distributed, the goodness of fit statistic ( $Q$ ) is reported as 0.533, which is considered fairly good, as it suggests that the remaining model residuals are likely a factor of random measurement error. The obtained activation energy  $E_a$  and reference rate constant  $k_{ref}$  (at  $T_{ref} = 33\text{ }^\circ\text{C}$ ) are listed in Table 1. Those parameters are comparable to those for acetyl

**Table 1. Fitting Results of the Kinetic Model**

	final value	units	standard error	confidence interval (%)
$k_{T_{ref} = 33\text{ }^\circ\text{C}}$	0.0721	$\text{h}^{-1}$	0.00019	3.32
$E_a$	107.2	$\text{kJ/mol}$	1.013	1.89

phosphate hydrolysis reported previously.<sup>12</sup> It is worth mentioning that the model established in this study is based on the assumption that pseudo-first-order reaction kinetics holds true at near-neutral pH, as reported previously.<sup>10</sup> The pH monitored during dynamic experiments is between 6 and 8.5 (Figure S10). For other reaction systems with pH-dependent kinetics, the designed temperature ramping may change the system pH and reaction mechanism, making method re-evaluation necessary.

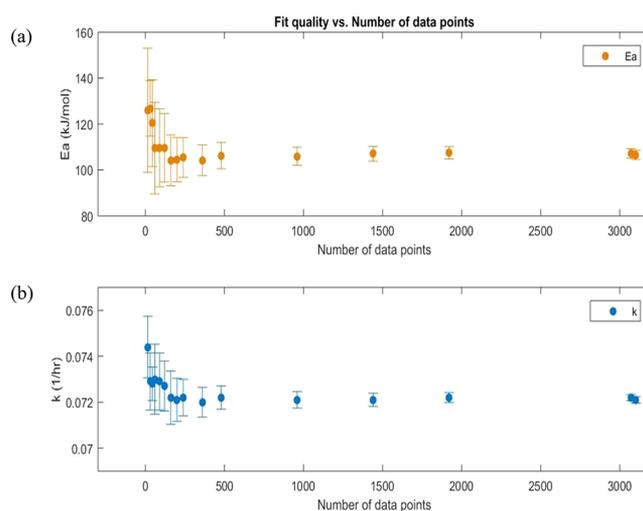
**3.2.2. Effect of Number of Data Points on the Estimation of the Reaction Kinetics.** With the application of in situ PAT and lab automation, such as with the ReactIR system employed in these experiments, it is possible to collect enormous datasets with several thousand data points per experiment. Although this technology represents a tremendous advancement for the collection and analysis of kinetic data, it presents a challenge for model-based regression of kinetic parameters in software such as DynoChem. Since the solver is required to accurately evaluate the system of equations at each data point to evaluate the model residuals in a regression, several thousand data points can lead to even more numerous model iterations, which can severely increase the processing time required to obtain a suitable parameter fit.

The quality of the parameter models can be preserved with a potentially smaller number of data points using appropriate data reduction algorithms. For this work, the large FTIR composition data sets were reduced to various sizes using the algorithm reported by Visvalingam and Whyatt.<sup>31</sup> This algorithm is often used for the reduction of visual and graphical data, such as for the compression of maps and images,<sup>32</sup> by nature of its design to preserve the appearance of complex shapes with the minimal number of required data points. In a similar way, in this work the algorithm was used to reduce the amount of kinetic profile data while preserving the key features of the curvature.

In brief, the algorithm of Visvalingam and Whyatt removes points from a series of data by considering the area of each triangle created by three consecutive points in the series. The point that lies in the center of the triangle of the lowest area is removed, and the triangle areas are recalculated from the remaining points. Points that create larger triangles, such as

where the response is changing significantly, are generally kept. On the other hand, a point without change that lies very close to its two neighbors will likely be removed. This procedure is followed iteratively until the desired number of data points remains. What we find is that this data reduction method preserves important features of our kinetic data where there is a steep slope or irregular/discontinuous shape. We expect that these regions of our data are of the highest importance to our model since they are expected to correlate with regions of high reactivity or changes in reactivity, which we would expect to confer high parameter sensitivity for regressions of good quality.

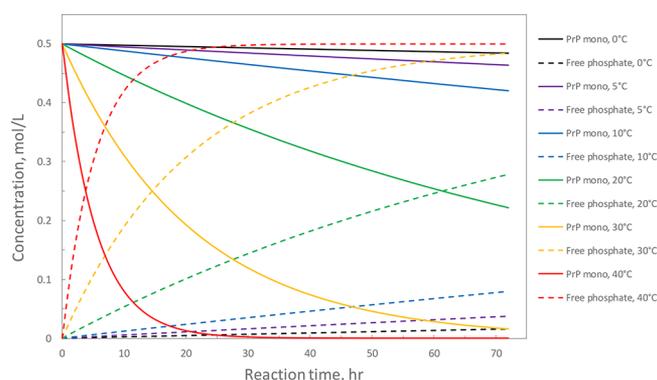
Figure 7 displays the estimated kinetic parameters  $E_a$  and  $k_{ref}$  as functions of the number of data points used for kinetic



**Figure 7.** Values of the kinetic parameters (a)  $E_a$  and (b)  $k_{ref}$  (at  $T_{ref} = 33\text{ }^\circ\text{C}$ ) estimated by DynoChem as functions of the number of data points.

model fitting. Evidently, consistent parameter estimations were obtained when at least 250 data points were included in the calculation. Such a large number of data points is extremely difficult, if not impossible, to acquire using a traditional off-line analytical method such as HPLC. Especially for cases like hydrolysis reactions, which continue to proceed even during HPLC off-line analysis under aqueous conditions, the concentration determined may be deemed inaccurate. On the contrary, PAT can readily generate sufficient data in situ with minimum effort spent on sample preparation and analysis without concern for sample stability. It is a powerful tool for gaining process understanding in an efficient and cost-effective way.

**3.2.3. Simulation of Reaction Performance.** As we are confident about the parameters obtained from the reaction performance analysis and modeling fitting, the degradation/hydrolysis reactions at various temperatures were simulated accordingly to help make predictions and further generate an engineering control strategy for risk mitigation. The simulation results in Figure 8 show that only 1% PrP mono was degraded at  $0\text{ }^\circ\text{C}$  within 24 h, while contrarily most of the PrP mono (98.7%) was degraded at  $40\text{ }^\circ\text{C}$  over the same period of time. In the desired enzymatic ATP regeneration reaction system, propionyl phosphate in this study is expected to be utilized as the phosphorylating agent in the presence of the enzyme acetate kinase to transfer phosphate group into ATP, which is



**Figure 8.** Concentration profiles of PrP mono (solid lines) and free phosphate (dashed lines) predicted by the DynoChem model at different temperatures.

further utilized as the target substrate to form the product. However, at elevated temperature, the PrP hydrolysis side reaction largely competes with the target enzymatic reaction. From this kinetics modeling fitting/simulation study, we learn that proper system temperature control (e.g.,  $<4\text{ }^{\circ}\text{C}$ ) should be put in place during and after PrP mono is charged to the vessel in the scale-up process in order to minimize undesired hydrolysis side reaction before other key raw materials and enzymes are introduced to trigger the target reaction.

#### 4. CONCLUSION

In this study, the hydrolysis reaction of propionyl phosphate monoammonium salt was carried out with a modified repeated temperature scanning experiment and closely monitored by real-time IR spectroscopy calibrated to off-line NMR analysis to obtain rich concentration data assisted by multivariate IR modeling. The concentration and reaction temperature profiles were fit to a first-order kinetic model with valid statistical indications by the DynoChem model, reporting two key kinetics parameters of PrP hydrolysis for the first time. The activation energy of PrP mono hydrolysis at near-neutral pH was evaluated to be 107.2 kJ/mol, and apparent rate constant at  $33\text{ }^{\circ}\text{C}$  ( $k_{T_{ref}=33\text{ }^{\circ}\text{C}}$ ) was  $0.0721\text{ h}^{-1}$ . The work presented herein demonstrates how the modified repeated temperature scanning method combined with PAT in situ monitoring of reaction performance can aid in efficiently producing quantifiable reaction kinetics and process understanding. This methodology allows scientists to use limited amounts of expensive pharmaceutical intermediates and constrained time and human resources to perform data-rich experiments for building models and guiding scale-up control strategy. We also anticipate that this method can be utilized for kinetic exploration of similar first-order homogeneous reactions or even other more complex reactions in synthetic chemistry.

#### ■ ASSOCIATED CONTENT

##### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.oprd.0c00451>.

Experimental procedure of quantitative  $^{31}\text{P}$  NMR spectroscopy,  $^{31}\text{P}$  NMR data analysis for concentration of starting material and product,  $^{31}\text{P}$  NMR spectra of all in-process samples, additional dynamic experiment for model validation (experimental procedure and results), and pH profile of the dynamic experimental run (PDF)

#### ■ AUTHOR INFORMATION

##### Corresponding Author

**Cuixian Yang** – Small Molecular Process Research & Development (SMPRD), MRL, Merck & Co., Inc., Rahway, New Jersey 07065, United States; [orcid.org/0000-0002-2470-0840](https://orcid.org/0000-0002-2470-0840); Phone: 732-594-6402; Email: [yang.cuixian@gmail.com](mailto:yang.cuixian@gmail.com)

##### Authors

**Hanzhou Feng** – Process Analytical Technology, MMD, Merck & Co., Inc., Rahway, New Jersey 07065, United States

**Kevin Stone** – Small Molecular Process Research & Development (SMPRD), MRL, Merck & Co., Inc., Rahway, New Jersey 07065, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.oprd.0c00451>

##### Notes

The authors declare no competing financial interest.

#### ■ ACKNOWLEDGMENTS

The authors thank Kevin R. Campos, Elizabeth S. Fisher, and Thomas P. Vickery (all at Merck & Co., Inc.) for providing insightful feedback on the manuscript.

#### ■ REFERENCES

- Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; Halsey, H. M.; Hughes, G. J.; Jo, J.; Joyce, L. A.; Kolev, J. N.; Liang, J.; Maloney, K. M.; Mann, B. F.; Marshall, N. M.; McLaughlin, M.; Moore, J. C.; Murphy, G. S.; Nawrat, C. C.; Nazor, J.; Novick, S.; Patel, N. R.; Rodriguez-Granillo, A.; Robaire, S. A.; Sherer, E. C.; Truppo, M. D.; Whittaker, A. M.; Verma, D.; Xiao, L.; Xu, Y.; Yang, H. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **2019**, *366*, 1255–1259.
- Iglesias, L. E.; Lewkowicz, E. S.; Medici, R.; Bianchi, P.; Iribarren, A. M. Biocatalytic approaches applied to the synthesis of nucleoside prodrugs. *Biotechnol. Adv.* **2015**, *33*, 412–434.
- Jordheim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat. Rev. Drug Discovery* **2013**, *12*, 447.
- Andexer, J. N.; Richter, M. Emerging enzymes for ATP regeneration in biocatalytic processes. *ChemBioChem* **2015**, *16*, 380–6.
- Langer, R. S.; Hamilton, B. K.; Gardner, C. R.; Archer, M. C.; Colton, C. K. Enzymatic regeneration of ATP. I. Alternative routes. *AIChE J.* **1976**, *22*, 1079–1090.
- Langer, R. S.; Gardner, C. R.; Hamilton, B. K.; Colton, C. K. Enzymatic regeneration of ATP: II. Equilibrium studies with acetate kinase and adenylate kinase. *AIChE J.* **1977**, *23*, 1–10.
- Baughn, R. L.; Adalsteinsson, O.; Whitesides, G. M. Large-scale enzyme-catalyzed synthesis of ATP from adenosine and acetyl phosphate. Regeneration of ATP from AMP. *J. Am. Chem. Soc.* **1978**, *100*, 304–306.
- Crans, D. C.; Kazlauskas, R. J.; Hirschbein, B. L.; Wong, C. H.; Abril, O.; Whitesides, G. M. Enzymatic regeneration of adenosine 5'-triphosphate: acetyl phosphate, phosphoenolpyruvate, methoxycarbonyl phosphate, dihydroxyacetone phosphate, 5-phospho-alpha-D-ribose 5-phosphate, uridine-5'-diphosphoglucose. *Methods Enzymol.* **1987**, *136*, 263–80.
- Kazlauskas, R. J.; Whitesides, G. M. Synthesis of methoxycarbonyl phosphate, new reagent having high phosphoryl donor potential for use in ATP cofactor regeneration. *J. Org. Chem.* **1985**, *50*, 1069–1076.

- (10) Koshland, D. E. Effect of Catalysts on the Hydrolysis of Acetyl Phosphate. Nucleophilic Displacement Mechanisms in Enzymatic Reactions. *J. Am. Chem. Soc.* **1952**, *74*, 2286–2292.
- (11) Bentley, R. The Mechanism of Hydrolysis of Acetyl Dihydrogen Phosphate. *J. Am. Chem. Soc.* **1949**, *71*, 2765–2767.
- (12) Oestreich, C. H.; Jones, M. M. The effect of metal ions on labile phosphates. I. The hydrolysis of acetyl phosphate dianion. *Biochemistry* **1966**, *5*, 2926–31.
- (13) Park, J. H.; Koshland, D. E., Jr. The hydrolytic activity of glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **1958**, *233*, 986–90.
- (14) *Guidance for Industry: PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*; U.S. Food and Drug Administration, 2004.
- (15) Arrondo, J. L. R.; Muga, A.; Castresana, J.; Goni, F. M. Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. *Prog. Biophys. Mol. Biol.* **1993**, *59*, 23–56.
- (16) Du, X.; Frei, H.; Kim, S. H. The Mechanism of GTP Hydrolysis by Ras Probed by Fourier Transform Infrared Spectroscopy. *J. Biol. Chem.* **2000**, *275*, 8492–8500.
- (17) Pacheco, R.; Karmali, A.; Serralheiro, M. L.; Haris, P. I. Application of Fourier transform infrared spectroscopy for monitoring hydrolysis and synthesis reactions catalyzed by a recombinant amidase. *Anal. Biochem.* **2005**, *346*, 49–58.
- (18) Flynn, J. H.; Wall, L. A. A quick, direct method for the determination of activation energy from thermogravimetric data. *J. Polym. Sci., Part B: Polym. Lett.* **1966**, *4*, 323–328.
- (19) Blaine, R. L.; Kissinger, H. E. Homer Kissinger and the Kissinger equation. *Thermochim. Acta* **2012**, *540*, 1–6.
- (20) Friedman, H. L. New methods for evaluating kinetic parameters from thermal analysis data. *J. Polym. Sci., Part B: Polym. Lett.* **1969**, *7*, 41–46.
- (21) Ozawa, T. Kinetic analysis by repeated temperature scanning. Part 1. *Thermochim. Acta* **2000**, *356*, 173–180.
- (22) Wojciechowski, B. W. The temperature scanning reactor: A kinetics instrument. *Chem. Eng. Commun.* **2003**, *190*, 1115–1131.
- (23) Schmidt, O. P.; Dechert-Schmitt, A.-M.; Garnsey, M. R.; Wisniewska, H. M.; Blackmond, D. G. Kinetic Analysis of Catalytic Organic Reactions Using a Temperature Scanning Protocol. *ChemCatChem* **2019**, *11*, 3808–3813.
- (24) Asprey, S. P.; Wojciechowski, B. W.; Rice, N. M.; Dorcas, A. Applications of temperature scanning in kinetic investigations: The hydrolysis of acetic anhydride. *Chem. Eng. Sci.* **1996**, *51*, 4681–4692.
- (25) McMullen, J. P.; Jensen, K. F. Rapid Determination of Reaction Kinetics with an Automated Microfluidic System. *Org. Process Res. Dev.* **2011**, *15*, 398–407.
- (26) Aroh, K. C.; Jensen, K. F. Efficient kinetic experiments in continuous flow microreactors. *React. Chem. Eng.* **2018**, *3*, 94–101.
- (27) Moltgen, C. V.; Puchert, T.; Menezes, J. C.; Lochmann, D.; Reich, G. A novel in-line NIR spectroscopy application for the monitoring of tablet film coating in an industrial scale process. *Talanta* **2012**, *92*, 26–37.
- (28) Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
- (29) More, J. J. *Levenberg–Marquardt Algorithm: Implementation and Theory*; Argonne National Laboratory: Argonne, IL, 1977.
- (30) Wilamowski, B. M.; Yu, H. Improved Computation for Levenberg–Marquardt Training. *IEEE Trans. Neural Networks* **2010**, *21*, 930–937.
- (31) Visvalingam, M.; Whyatt, J. D. Line generalisation by repeated elimination of points. *Cartogr. J.* **1993**, *30*, 46–51.
- (32) Visvalingam, M.; Williamson, P. J. Simplification and Generalization of Large Scale Data for Roads: A Comparison of Two Filtering Algorithms. *Cartogr. Geogr. Inf. Syst.* **1995**, *22*, 264–275.