Chemical Engineering Journal 523 (2025) 167770

Contents lists available at ScienceDirect

Chemical Engineering Journal

journal homepage: www.elsevier.com/locate/cej

Identifying a kinetic model for the acid-catalyzed sequence of the artemisinin partial synthesis

Moritz Schulze [a,b,c], Susann Triemer [d,e], René Schenkendorf [f], Andreas Seidel-Morgenstern [d,e], Ulrike Krewer [a,*]

[a] *Institute for Applied Materials – Electrochemical Technologies, Karlsruhe Institute of Technology, Adenauerring 20b, 76131, Karlsruhe, Germany*
[b] *Institute of Energy and Process Systems Engineering, Technische Universität Braunschweig, Langer Kamp 19B, 38106, Braunschweig, Germany*
[c] *Center of Pharmaceutical Engineering, Technische Universität Braunschweig, Franz-Liszt-Straße 35a, 38106, Braunschweig, Germany*
[d] *Institut für Verfahrenstechnik, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 39106, Magdeburg, Germany*
[e] *Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106, Magdeburg, Germany*
[f] *Automation & Computer Sciences Department, Harz University of Applied Sciences, Friedrichstraße 57-59, 38855, Wernigerode, Germany*

ARTICLE INFO

Dataset link: https://doi.org/10.35097/2vwx8arcenkf1jpc

Keywords:
Malaria
Acid-catalyzed reaction
Continuous manufacturing
Photo-flow reactor
Oxidation processes
Model selection

ABSTRACT

Artemisinin derivatives, recognized as the most potent antimalarial agents with extensive biological activity, can be produced via semi-synthetic conversion of dihydroartemisinic acid through a hydroperoxide intermediate (PO_1). However, modeling and optimizing this synthesis, particularly the acid-catalyzed transformation of PO_1 into artemisinin, has been challenging due to complex and partly unknown reaction pathways, and coupled mass-transfer processes. Extending our previous work, we present the first semi-empirical kinetic model describing the partial synthesis of artemisinin. Our developed identification approach integrates chemical domain expertise, mathematical optimization, and rigorous model evaluation against experimental data. Data were collected from a milliscale photo-flow and a batch reactor, leveraging both the efficient mass transfer of the continuous system and the frequent sampling capability of the batch setup. Numerical analysis revealed that the acid catalyst influences reaction kinetics more significantly than previously assumed. The identified kinetic model features apparent reaction orders greater than two with respect to the concentration of the added acid, and it contains different intermediates and byproducts depending on whether the operation is conducted in the presence of oxygen or under anaerobic conditions. Although the semi-empirical nature of the approach necessitates careful interpretation of these mechanistic conclusions, the developed process model accurately reproduces the behavior of the reaction system, providing parameter estimates crucial for process optimization.

1. Introduction

Drug therapies based on artemisinin (ART) show the highest efficacy against malaria and are therefore recommended as first-line treatment against the disease [1]. Next to its antimalarial properties, ART and its derivatives exhibit further biological activity, e.g., anticancer, antibacterial or antiviral [2], and are thus an active research field [3–5]. ART is industrially supplied exclusively by extraction from the leaves of the plant *Artemisia annua*. Alternatively, semi-synthetic routes for artemisinin production have been targeted to be industrialized [6–8]; these are based on the fermentative production of ART precursors. The synthesis steps following heterologous production to ART have in common that dihydroartemisinic acid (DHAA) is converted to ART in a two-step process based on the fundamental works by Acton and Roth [9–11]: A photooxygenation of DHAA yielding a tertiary hydroperoxide (PO_1) is followed by an acid-catalyzed reaction cascade that converts PO_1 to ART, Fig. 1. In contrast to the industrial batch approaches, continuous-flow photochemical conversion [12,13] of DHAA to ART has been advancing [14] and a corresponding process was introduced by Lévesque and Seeberger [15]. The developed two-phase flow reactor showed a very short residence time of $11.5\,\mathrm{min}$, an ART yield of 65% and a space–time yield of $3.5\,\mathrm{g\,mL^{-1}\,d^{-1}}$ [16]. This approach was further extended to utilize the DHAA content of the plant itself, and made use of the naturally present chlorophylls in the extract as photosensitizers for a green chemical synthesis [17].

Despite the published research on the conversion of DHAA to ART, the detailed elucidation of the reaction mechanism is still being investigated [18–20]. However, for a reliable and safe process design to cost-efficiently produce the drug molecule, the underlying phenomena, in particular the mechanism and kinetics of the reaction network, have to be well understood. To identify chemical reaction networks and their

---

[*] Corresponding author.
E-mail address: ulrike.krewer@kit.edu (U. Krewer).

https://doi.org/10.1016/j.cej.2025.167770
Received 9 February 2025; Received in revised form 8 August 2025; Accepted 26 August 2025
Available online 2 September 2025
1385-8947/

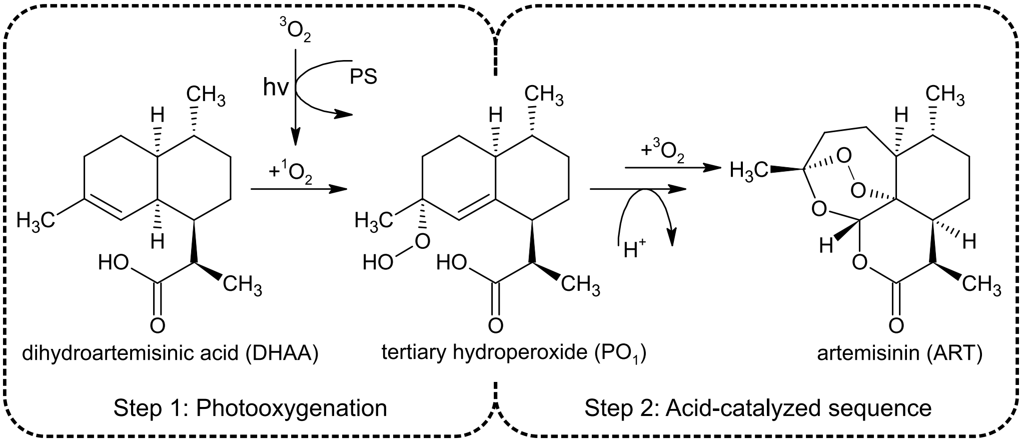


Fig. 1. Simplified semi-synthetic path from dihydroartemisinic acid (DHAA) to artemisinin (ART) based on Acton and Roth [9–11] with the singlet oxygen formed by a photo-induced process (PS: photosensitizer).

dynamics in complex synthesis processes for finally designing the process, mathematical modeling provides a powerful tool [21,22]. In [23], the authors of this article published a process model for the photooxygenation step of the artemisinin partial synthesis from DHAA to PO_1 in a two-phase milliscale photo-flow reactor. The here presented work enhances the previous research and aims at identifying a semi-empirical kinetic model for the second step of the partial synthesis, the acid-catalyzed sequence from PO_1 to ART. To the best of our knowledge, no kinetic model for the complete reaction sequence to artemisinin has yet been developed. This study is therefore intended to provide a first kinetic description of the reactions leading to artemisinin based on a semi-empirical strategy. A full mechanistic understanding including the elucidation of individual reaction pathways, elementary steps, or microkinetic analysis is beyond the scope of this study. We examine the reaction kinetics of the acid-catalyzed sequence using two experimental methods: continuous operation in the presence of oxygen and batch operation under a nitrogen atmosphere. A significant challenge hereby is the incomplete knowledge regarding the chemical reaction network (CRN) and the need to develop an engineering-oriented model rather than elucidating intrinsic kinetics.

Both aspects are common issues in the field of chemical engineering. Unraveling the mechanism and kinetics often follows from collected time-resolved species data [24–27]. Applying these tools to the acid-catalyzed reaction sequence is complicated, as the quantifiable product spectrum is incomplete and both data sets from the batch and continuous experiments cover different subsets of the CRN. Moreover, the experimental data from the flow reactor is collected when the reactor system has reached a steady state [28]. Therefore, methods based on time-resolved data were ruled out, and a knowledge-based semi-empirical strategy was chosen instead. This strategy integrates chemical mechanistic knowledge with mathematical optimization and statistical model selection via the Akaike information criterion [29]. For the underlying process modeling, simplified reactor descriptions were employed to enable identification of an initial dynamic process model for the acid-catalyzed sequence based on established literature mechanisms. This work thus presents a semi-mechanistic process model developed for the reactor configuration studied, with effective kinetic parameters. While not determining intrinsic kinetic constants, the model provides a reasonable framework for characterizing and optimizing the acid-catalyzed sequence of the artemisinin partial synthesis process.

Following this introductory section, the reaction mechanism based on published literature is discussed. Afterwards, the experimental setups, the kinetic modeling approach and the model identification strategy are explained. The results part is divided into an experimental data analysis and the identification of the final kinetic model candidate.

2. Reaction mechanism

The reaction mechanism of the acid-catalyzed conversion of the intermediate PO_1 to artemisinin was studied intensively to understand artemisinin formation within its natural source *Artemisia annua* and to boost partial-synthetic artemisinin production. This research began with Acton and Roth in 1992 and has since been expanded by contributions from Haynes and Vonwiller et al. [30–32], by Brown and Sy et al. [33–36], and by Arman, Varela and Yoshimoto [18–20]. The essential results have been used to formulate the simplified reaction scheme in Fig. 2. The main steps considered in the model identification of the acid-catalyzed sequence, along with the experimental rationale underpinning these steps, are discussed in the following. When adding a strong acid to the intermediate hydroperoxide PO_1, PO_1 is converted via a Hock-cleavage [10] and a nucleophilic attack of water to an unstable seven-membered hemiacetal. However, this postulated intermediate could not be identified in experiments so far [31,35]. The hemiacetal is again cleaved in acid-catalyzed steps to a semi-stable enol, that was managed to be isolated and analyzed at temperatures below −20 °C [31,32,35]. Besides the reaction pathway to the enol, PO_1 can be converted in an acid-catalyzed loss reaction of PO_1 to dihydro-epi-deoxyarteannuin B [35,36], which is in the model referred to as BP_3.

In the presence of triplet oxygen, the enol reacts further to a hydroperoxyl aldehyde that is subsequently rearranged acid-catalyzed to a peroxy-hemiacetal. Both intermediates have been isolated when a DHAA ester was used as starting material [30,31,35]. In the final step, the peroxy-hemiacetal converts to ART in a series of irreversible reactions and under the influence of a strong Brønstedt acid [30]. As a side reaction, the enol reacts to a seco-cadinane intermediate yielding the lactone BP_4 [31,35].

The proposed reaction mechanism along the hemiacetal and the enol is backed by experiments using labeled compounds [10,19,20]. The main products of the semi-synthesis from DHAA to ART are the target ART itself, and the following unwanted byproducts: arteannuin H, dihydro-epi-deoxyarteannuin B (BP_3), and the lactone BP_4.

Up to 80 % of used DHAA has been recovered in literature [10,16]. Many other reaction species have been identified or postulated [35], but a complete picture of the whole reaction mechanism is unknown as in all studies (substantial) amounts of converted DHAA could not be recovered in identified products. Despite the complex mechanism, high yields up to 71 % for the partial synthesis can be achieved [37]. Current research directions involve in particular the role of the solvent, the acidic catalyst, and the participating water [38,39].

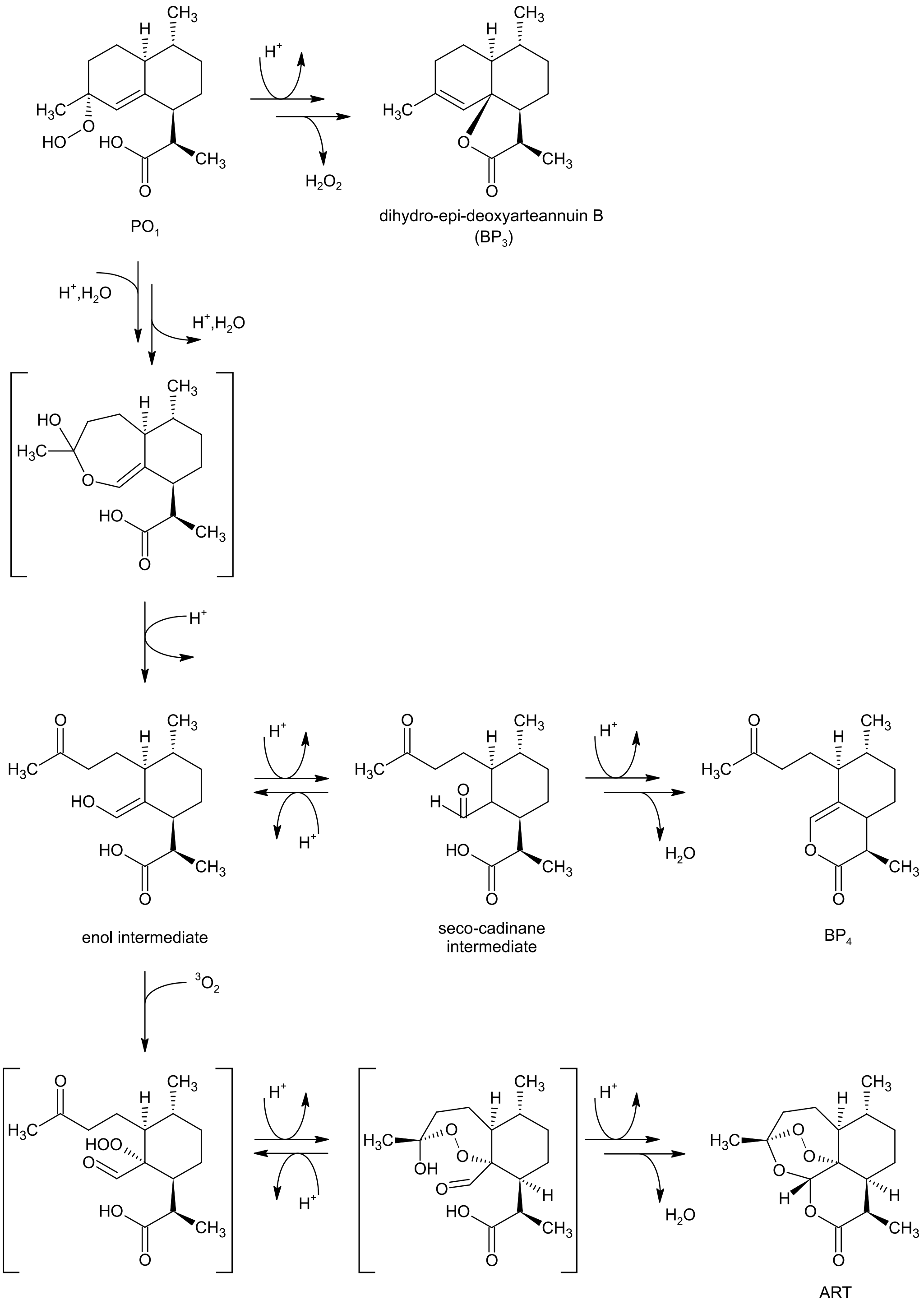


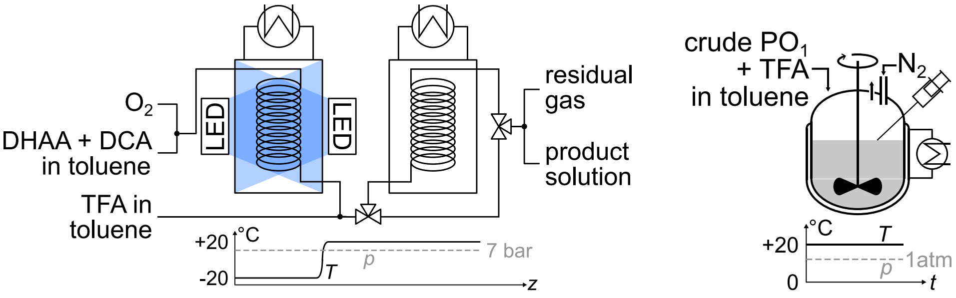
Fig. 2. Proposed reaction mechanism for the acid-catalyzed sequence forming ART from PO_1 [35]. Reaction with H^+ entering and leaving indicate acid-catalyzed steps. Various versions of kinetic models for this mechanism are proposed and analyzed in this work, see Figs. 8 and 10 and Appendix D.

3. Material and methods

3.1. Experimental setups and operating conditions

The conversion of the intermediate hydroperoxide (PO_1) to artemisinin was carried out using toluene as the solvent and trifluoroacetic acid (TFA) as the acidic catalyst. Both components exhibited effective performance in the overall partial synthesis, resulting in high yields of artemisinin [16]. Toluene was also employed in the preliminary investigation of photooxygenation kinetics [23]. Thus, the selection of these materials aligns with both the previous and current study.

The experimental data was obtained from the targeted continuous and a batch setup. Both differ in the oxygen availability, and the frequency samples could be taken. The continuous two-phase milliscale reactor setup (Fig. 3(a)) ensured high oxygen transfer rates, whereas the batch setup operated in an oxygen-free nitrogen atmosphere (Fig. 3(b)). The reason for conducting also batch experiments was to obtain better insights into the mechanism and kinetics by time-resolved experiments as the continuous setup was operated in steady-state mode



(a) (b)

Fig. 3. Experimental setups and operating conditions for the investigation of the acid-catalyzed sequence from the intermediate PO_1 to artemisinin: (a) Continuous setup to investigate the reaction in the presence of oxygen and (b) batch setup to enable time-resolved sampling.

and samples were only drawn at the end of the reaction line. The residence time range was limited to approx. 15 min due to the limit in the available reactor volume and the resulting duration of the system to reach a steady-state. The batch experiments allowed us to observe the conversion of PO_1 over a broader time range and, thus, enabled lower catalyst concentration of TFA in the acid-catalyzed experiments.

For the continuous experiments, the milliscale photo-flow-reactor introduced in our previous study [23] was complemented by an additional acid-dosing pump and a second reaction line consisting of FEP tubing (ID: 1.6 mm, OD: 3.2 mm) offering reaction volume for the reaction to artemisinin. This reaction line is wrapped around a stainless steel frame and integrated into a second stainless steel container connected to a second thermostat for temperature control. The length of the second reaction line was varied between 2 m and 6 m depending on the target residence time.

To separate the photooxygenation of DHAA to PO_1 and the acid-catalyzed sequence of PO_1 to ART, the acidic feed was dosed separately from the reactant containing feed. The reactant feed consisted of DHAA and the photosensitizer dicyanoanthracene (DCA) in toluene. The reactant feed was mixed with the oxygen feed (x_{O_2}=1) and pumped through the photoreactor kept at −20 °C. The residence time in the photoreactor was chosen to reach full DHAA conversion equal to a PO_1 yield of approx. 85 %. A second feed consisting of TFA in toluene was dosed to the reaction stream exiting the photoreactor and before entering the second reaction line initiating the acid-catalyzing solution. The ratio of the DHAA feed and the TFA feed was kept in all experiments at 9:1 (v/v). The TFA concentration in the acidic feed and the DHAA and DCA concentration in the reactant feed was set according to the target initial PO_1 and TFA concentration in the reaction solution. All experiments in the continuous setup were operated with a ratio of the gas to the total liquid feed of 4:1 (v/v). The total flow rate was varied in a tight range between 0.5 $\mathrm{mL\,min^{-1}}$ to 2.5 $\mathrm{mL\,min^{-1}}$ to ensure comparable Taylor flow conditions. The temperature in the acid-catalyzed section was set to 20 °C. The pressure in the entire reaction line was 7 bar. Samples of the reactor effluent were collected when the whole system was stable for at least 5 times of the theoretical residence time in the entire system, to ensure steady-state conditions. The system was defined as stable when the temperature in all reactors did not vary more than 0.1 °C and the pressure measured at the beginning and the end of the whole reactor line was kept between 6.8 bar bar to 7.2 bar. At each set of reaction conditions two samples were collected at the reactor outlet after phase separation and depressurization and quenched directly with triethylamine (TEA) to stop the reaction. Further details on the continuous setup and its operation can be found in former publications by the authors [23,28].

For the batch experiments, the reactant solution was prepared on the day before in the continuous setup. The feed containing 0.25 M DHAA and 2.5 mM DCA in toluene pumped through the photoreactor kept at −20 °C together with pure oxygen with a total flow rate of 2 $\mathrm{mL\,min^{-1}}$ so that a residence time of 2.5 min in the irradiated section was obtained. The reactor effluent was collected after depressurization in a brown flask and stored overnight in a fridge at 4 °C. The obtained solution was used – similarly as for the continuous setup – directly as the feed solution for the batch experiments without further workup. Using this procedure, an initial concentration of 0.2 M PO_1 was obtained.

The batch setup, Fig. 3(b), consisted of a double-walled 100 mL glass reactor mixed by a magnetic stirrer and attached to a cooling thermostat keeping the temperature within the reactor at 20 °C. The reactor was closed with a rubber stopper including one inlet and one outlet tubing for supplying nitrogen. At the start of the experiment, 30 mL of the PO_1 containing feed were stirred for 30 min while rinsed with nitrogen to remove all oxygen from the flask. The reaction was initiated by adding 1 mL toluene with dissolved TFA. The samples were drawn with a syringe after 1, 2, 3, 5, 7.5, 10, 15, 30, 60, and 90 min and quenched with TEA dissolved in 0.4 mL toluene. An overview of the constant and varied reaction conditions of both continuous and batch experiments is included in Appendix A.

All samples were analyzed by ^{1}H−NMR [28]. The analyzed measurement data was then used in an experimental data analysis of the synthesis process, Section 4.1, for which the following key measures for chemical reactions were evaluated.

3.2. Evaluating the performance of reactions

A superficial residence time can be calculated for a two-phase flow reactor from

$$\tau^{s} = \frac{V}{\dot{V}_{l,0} + \dot{V}_{g,0}}, \tag{1}$$

with V the considered reactor volume and $\dot{V}_{l,0}$ and $\dot{V}_{g,0}$ the initial liquid and gas volumetric flow, respectively.

The yield of the formed species C relative to the consumed species B in a constant volume reactor is [40]

$$Y_{C} = \frac{[C] - [C]_0}{[B]_0} \cdot \frac{-\nu_B}{\nu_C}, \tag{2}$$

where square brackets indicate concentration and the subscript 0 marks the reference state, e.g., the initial conditions. ν_B and ν_C are the stoichiometric numbers of the reactant and product, respectively, in a chemical reaction. The corresponding conversion of species B is [40]

$$X_{B} = \frac{[B]_0 - [B]}{[B]_0}. \tag{3}$$

Lastly, the recovery of the species B indicates the amount of species C_i descendant of B that could be found in measurements:

$$R_{B} = \frac{[B] + \sum_i [C_i] \frac{-\nu_{B,i}}{\nu_{C,i}}}{[B]_0}. \tag{4}$$

3.3. Kinetic and process modeling

To describe rates of chemical reactions in a kinetic model, the law of mass action is commonly applied. The rate of a chemical reaction is here expressed as [41]

$$r = k \prod_i c_i^{m_i} \tag{5}$$

with kinetic rate coefficient k, the concentrations of the participating reagents c_i and the exponents m_i referred to as reaction orders. From reaction rates, the net rate of reaction for a species i can be assembled:

$$r^i = \sum_j \nu_{i,j} r_j \tag{6}$$

with $\nu_{i,j}$ the stoichiometric number of species i in reaction j and r_j the corresponding rates of reaction from Eq. (5).

In non-elementary reactions, the reaction orders do not equal the reagents' stoichiometric coefficients as in their elementary counterparts [41]. However, non-elementary reactions can be seen as a combination of elementary reactions where intermediates form that occur in very low quantity and are therefore difficult to detect [41]. Similarly as for the rate coefficient k, the reaction order m in non-elementary reactions can be estimated from measurement data. It should be noted that the reaction orders m_i in Eq. (5) represent apparent (or effective) reaction orders specific to our reactor system. While our analysis suggests that transport limitations are minimized under the studied operating conditions, minor transport influences cannot be completely ruled out. The reaction orders, along with the reaction rate coefficients, should therefore not be interpreted as intrinsic kinetic parameters.

For deriving a process model, the kinetic model for simulating the chemical reaction network must be integrated in a reactor model that describes the experimental setup [42,43].

3.4. Process model for the batch setup

In the batch setup operating in oxygen-free atmosphere, a batch reactor model with constant reactor volume is applied [41]. Integrating kinetic rate laws, Eq. (6), the governing equations of the process model for the batch setup are

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = r^i = \sum_j \nu_{i,j} r_j \tag{7}$$

for all participating species i and identified reactions j and their rate laws. The corresponding initial conditions are

$$c_i(t=0) = \begin{cases} [\mathrm{PO}_1]_0 & \text{if } i = \mathrm{PO}_1, \\ [\mathrm{BP}_3]_0 & \text{if } i = \mathrm{BP}_3, \\ 0 & \text{else,} \end{cases} \tag{8}$$

where only PO_1 and BP_3 have non-zero initial conditions, as they are components of the prepared reactant solution, see Section 3.1.

3.5. Process model for the continuous setup

In the continuous setup, the reactor model from our previous publication is used [23]. It was developed for the same experimental setup and applied to the first step in the artemisinin partial synthesis, the photooxygenation (step 1 in Fig. 1). While the first reactor governs the light-induced reactions at −20 °C identified in our previous work, the second reactor involves the acid-catalyzed reaction network at 20 °C triggered by the addition of TFA downstream of the photo reactor, Fig. 3(a). The intermediate hydroperoxide PO_1 formed in the photoreactor is converted to artemisinin in the second reactor. Since the two steps describe different chemical processes, this work follows a sequential model development strategy to individually characterize the kinetics in the two reactors.

Simulated concentrations of the hydroperoxide of interest PO_1 in the photoreactor that are used in this work as an input to identify the kinetic model for the second step resulted in an average deviation of 7.1 % when compared to the measured concentrations [23]. The common reactor model is based on the drift-flux model [44], with key assumptions previously established [23] and summarized as follows. A discussion on the generalizability of the identified kinetic model under the consideration of the assumptions is given in Section 4.7.

- The conditioned gas and liquid flow rates establish Taylor flow conditions at all times. The total flow rate was varied over a narrow range of only $0.5\,\mathrm{mL\,min^{-1}}$ to $2.5\,\mathrm{mL\,min^{-1}}$ ensuring consistent hydrodynamic conditions across all experimental conditions. The experimentally determined distribution factor was 1.020 ± 0.025 (95 % confidence interval) [23] suggesting a very low slip between the two phases. Thus, the two phases' dynamics are closely coupled moving with similar velocities – an important feature of the drift-flux model [44]. Moreover, the high length-to-diameter ratios (10 m photoreactor length and 2 m to 6 m synthesis reactor length with 0.8 mm and 1.6 mm diameter tubing, respectively), enable continuum representations as employed in drift flux modeling.
- The mass transfer between gas and liquid phase is modeled using the usually applied linear approach [45] and is based on [46]. The authors analyzed mass transfer of oxygen into water under Taylor flow in circular cross-sections with inner diameters ranging from 1 to 3 mm. With the film contribution being the primary factor, the mass transfer coefficient could be described as $k_l a \propto \sqrt{\frac{D u_\mathrm{g}^\mathrm{s}}{L_\mathrm{UC}}}\frac{1}{d}$ with D the liquid diffusivity, u_g^s the superficial gas velocity, L_UC the length of a unit cell, and d the inner diameter of the tube. Assuming low fluctuations in liquid diffusivity and that the effect of varying superficial gas velocities across different experiments ($0.4\,\mathrm{mL\,min^{-1}}$ to $2\,\mathrm{mL\,min^{-1}}$ initial gas flow rates at 20 °C) dominates over changes in the unit cell length along the reaction line, the mass transfer coefficient is reduced to $k_l a = \widetilde{k_l a}\sqrt{u_\mathrm{g}^\mathrm{s}}$. The constant $\widetilde{k_l a}$ thus integrates the effect of the unit cell length and the toluene-gas system.
- The setup was operated in steady-state. No transient terms are therefore considered during modeling.
- A momentum balance is not considered as the relative pressure drop over the reactor is small (0.1 bar to 0.5 bar at 7 bar operating pressure).
- Energy balances are neglected due to the isothermal operation of the reactor.
- The flow is one-dimensional (z axis). Hence, ideal mixing in radial direction is assumed.
- The gas phase can be described by an ideal gas mixture, and the liquid phase is considered incompressible.
- The saturated dissolved oxygen concentration is derived from Henry's law [23]. Corresponding parameters for oxygen in toluene are taken from [47]. The authors collected data in the ranges from 1 to 10 bar and from 20 to 110 °C:

  $$\ln(H) = \frac{A}{T} + B, \quad A = 400.426, \; B = 3.393, \tag{9}$$

  with T the temperature in Kelvin and H the molar fraction based Henry constant in MPa.

A schematic representation of the continuous reactor system as described by the developed mathematical process model is shown in Fig. 4. In the drift flux model, the gas fraction α constitutes the ratio of the cross-sectional area covered by the gas flow A_g to the total cross-sectional channel area A:

$$\alpha = \frac{A_\mathrm{g}}{A}. \tag{10}$$

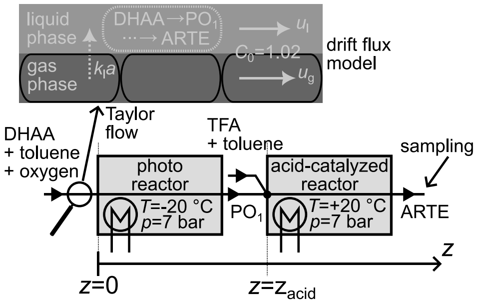


Fig. 4. Schematic representation of the 1-dimensional process model of the continuous two-phase reactor system, Fig. 3(a). The setup is operated in steady-state Taylor flow mode at constant pressure p and at constant, reactor-dependent temperatures; concentrations change spatially in flow direction z. The experimentally derived distribution parameter C_0 in the drift flux model causes a small deviation between the velocities of the liquid (u_{l}) and the gas phase (u_{g}). Mass transfer of oxygen occurs from the gas to the liquid phase ($k_{\mathrm{l}}a$). Reactions happen in the liquid phase with the light-induced reactions in the photo reactor and the acid-catalyzed reaction sequence in the second reactor.

Thus, the cross-sectional area covered by the liquid flow is $A_{\mathrm{l}}=(1-\alpha)A$. A constitutive equation for the phases' interaction relates the gas fraction

$$\alpha = \frac{1}{C_0}\beta, \tag{11}$$

to the known gas holdup β defined by the volumetric phase flows $\dot{V}_{\mathrm{g}}$ (gas) and $\dot{V}_{\mathrm{l}}$ (liquid):

$$\beta = \frac{\dot{V}_{\mathrm{g}}}{\dot{V}} = \frac{\dot{V}_{\mathrm{g}}}{\dot{V}_{\mathrm{g}} + \dot{V}_{\mathrm{l}}}, \tag{12}$$

where $C_0 = 1.02$ is the experimentally determined distribution factor from our previous study [23]. The material balance of a species i in the liquid phase becomes

$$\frac{\mathrm{d}c_i}{\mathrm{d}z} = \underbrace{\frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}}_{1/u_{\mathrm{l}}}\left(r^i + \delta_i j_{\mathrm{O_2}}\right), \quad \delta_i = \begin{cases} 1, & i = \mathrm{O_2} \\ 0, & \text{else,} \end{cases} \tag{13}$$

where u_{l} is the liquid phase velocity and r^i is the rate of concentration change of the considered species, Eq. (6). The mass transfer of oxygen from the gas to the liquid phase is

$$j_{\mathrm{O_2}} = \widetilde{k_{\mathrm{l}}a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}\left([\mathrm{O_2}]^\infty - [\mathrm{O_2}]\right) \tag{14}$$

with $[\mathrm{O_2}]^\infty$ the saturation concentration of oxygen in the liquid phase.

Analogously for the gas phase, the material balance over the total gas flow yields

$$\frac{\mathrm{d}\dot{V}_{\mathrm{g}}}{\mathrm{d}z} = -\frac{\bar{R}T}{p}(1-\alpha)Aj_{\mathrm{O_2}} \tag{15}$$

with $\bar{R}$ the universal gas constant, T the temperature and p the total pressure.

The reaction network and the governing equations of the photoreactor are given in Appendix C and were derived in our previous publication [23]. The focus of this publication is the reactor with the acid-catalyzed reactions. Its corresponding governing equations are

$$\begin{aligned}
\frac{\mathrm{d}[\mathrm{PO_1}]}{\mathrm{d}z} &= \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}\left(r_{\mathrm{acid}}^{\mathrm{PO_1}} - k_{\mathrm{PO_x}}[\mathrm{PO_1}]\right),\\
\frac{\mathrm{d}[\mathrm{PO_y}]}{\mathrm{d}z} &= \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}\left(-k_{\mathrm{PO_x}}[\mathrm{PO_y}]\right),\\
\frac{\mathrm{d}[\mathrm{PO_x}]}{\mathrm{d}z} &= \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}k_{\mathrm{PO_x}}\left([\mathrm{PO_1}] + [\mathrm{PO_y}]\right),\\
\frac{\mathrm{d}c_i}{\mathrm{d}z} &= \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}r_{\mathrm{acid}}^{i},\\
\frac{\mathrm{d}[\mathrm{O_2}]}{\mathrm{d}z} &= \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_{\mathrm{l}}}\left(r_{\mathrm{acid}}^{\mathrm{O_2}} + \widetilde{k_{\mathrm{l}}a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}([\mathrm{O_2}]^\infty - [\mathrm{O_2}])\right),\\
\frac{\mathrm{d}\dot{V}_{\mathrm{g}}}{\mathrm{d}z} &= -\frac{T}{p}\bar{R}\frac{\beta}{C_0}A\widetilde{k_{\mathrm{l}}a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}([\mathrm{O_2}]^\infty - [\mathrm{O_2}]).
\end{aligned} \tag{16}$$

r_{acid}^{i} are the net rates of reaction for species i of the acid-catalyzed reaction network that will be identified in the remainder of this article. Analogously, the differential equation for c_i represents a placeholder for each of the species taking part in the reaction network. $\mathrm{PO_y}$ are hydroperoxides formed in parallel with $\mathrm{PO_1}$ in the photoreactor. $\mathrm{PO_x}$ is a lumped loss product from the hydroperoxides $\mathrm{PO_1}$ and $\mathrm{PO_y}$ with $k_{\mathrm{PO_x}}$ the corresponding rate constant. The corresponding initial conditions at the inlet of the acid-catalyzed reactor at z_{acid} are

$$\begin{aligned}
&([\mathrm{PO_1}], [\mathrm{PO_y}], [\mathrm{PO_x}], c_i, [\mathrm{O_2}], \dot{V}_{\mathrm{g}})(z = z_{\mathrm{acid,0}})\\
&= ([\mathrm{PO_1}]_{\mathrm{acid,0}}, [\mathrm{PO_y}]_{\mathrm{acid,0}}, [\mathrm{PO_x}]_{\mathrm{acid,0}}, 0, [\mathrm{O_2}]_{\mathrm{acid,0}}, \dot{V}_{\mathrm{g,acid,0}}),
\end{aligned} \tag{17}$$

where the non-zero conditions follow from simulating the upstream part of the setup including the photoreactor, Fig. 4. The only non-zero conditions for the photoreactor at its inlet are $[\mathrm{DHAA}](z{=}0){=}[\mathrm{DHAA}]_0$, $[\mathrm{O_2}](z{=}0){=}[\mathrm{O_2}]^\infty$ and $\dot{V}_{\mathrm{g}}(z{=}0){=}\dot{V}_{\mathrm{g,0}}$ with the initial DHAA concentration and the initial $\dot{V}_{\mathrm{g}}$ being experimental settings. Note that the liquid flow rates at the photoreactor inlet (DHAA feed) and at the acid-catalyzed reactor (TFA feed), Fig. 4, are predetermined by fixed ratios, see Section 3.1. The reaction network of the photoreactor and its governing equations are given in Appendix D.

3.6. Estimating model parameters

Both process models for the batch and the continuous setup are mathematically described by a system of ordinary differential equations. The unknown model parameters $\boldsymbol{\theta}$ of the process model to be estimated are

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{k} \\ \boldsymbol{m} \end{pmatrix} \in \mathbb{R}^q, \tag{18}$$

where the vector of kinetic rate coefficients $\boldsymbol{k}$ and the vector of reaction orders $\boldsymbol{m}$ of the considered chemical reactions, Eq. (5), are concatenated. The elements of the two vectors and therefore the number of unknown model parameters q are specific for each kinetic model candidate. Note that for model candidates with fixed reaction orders, $\boldsymbol{m}$ is not present in $\boldsymbol{\theta}$.

With the exception of $k_{\mathrm{PO_x}}$, kinetic parameter values for the photooxygenation step were taken from [23]. The parameter $k_{\mathrm{PO_x}}$ is re-estimated in this study because the original parameterization accounted for hydroperoxide losses occurring between sampling and measurement. However, in the acid-catalyzed sequence, such losses are significantly reduced since the primary hydroperoxide $\mathrm{PO_1}$ reacts immediately upon acid addition.

An overview of all the model parameters transferred from the previous study and used here in the developed process model to identify the acid-catalyzed reaction network, is given in Table C.4 in Appendix A. Since model parameters for the photo reactor were taken from a previous study, their attached parametric uncertainty potentially

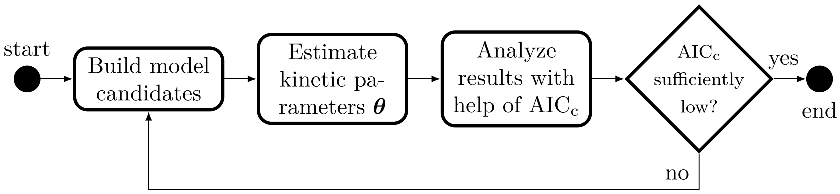


Fig. 5. Knowledge-based systematic strategy for kinetic model identification.

increases the uncertainty of the model parameter estimation for the acid-catalyzed sequence in this study. This effect will be negligible because a combined global sensitivity analysis of the estimated model parameters and the process settings in the photo reactor showed that the relative influence of the model parameters on the PO_1 concentration was vanishingly small [48].

For the mass transfer coefficient $\widetilde{k_1 a}$, parameter identifiability issues arose in the acid-catalyzed reactor system. The absence of gas phase sampling and liquid sampling only after oxygen equilibration made direct observation of mass transfer dynamics infeasible, rendering $\widetilde{k_1 a}$ non-identifiable from the collected data. Consequently, the coefficient was fixed at the value estimated for the photo reactor following an engineering compromise that is necessitated by the available data and system complexity constraints. A preliminary evaluation of this approximation can be made by examining the three competing factors affecting the mass transfer coefficient, i.e., $k_1 a = \widetilde{k_1 a}\sqrt{u_g^s} \propto \sqrt{\frac{D u_g^s}{L_{UC}}\frac{1}{d}}$. While the diameter doubles from the photo to the acid-catalyzed reactor, the 40 °C increased temperature causes the oxygen diffusivity to rise. For the comparable carbon dioxide molecule, the diffusivity in toluene increases by approximately 1.8 from −20 °C to 20 °C [49]. The most difficult factor to assess is the unit cell length L_{UC}. From a purely geometric perspective, the unit cell length must decrease by a factor of four when the diameter is doubled. However, the situation is significantly more complex due to hydrodynamic effects at the diameter change [50], making it impossible to provide a definitive estimate. Given these competing and complex effects and that mass transfer is not expected to be rate-limiting under the studied conditions, fixing $\widetilde{k_1 a}$ to the photo reactor value represents a reasonable engineering approximation.

3.7. Strategy for identifying kinetic models

The developed identification strategy builds upon an iterative cycle starting with the knowledge-based proposition of kinetic model candidates, displayed in Fig. 5. The parameters of the different model candidates are then estimated from the available experimental data. Supported by a model selection criterion, the subsequently introduced Akaike information criterion AIC$_c$, one final model candidate is chosen from the set by the expert. Alternatively, if none of the kinetic model candidates is sufficient, i.e., the AIC$_c$ value is too large or artefacts are observed in the model-data fit, further candidates are postulated and the procedure is repeated.

A well-known model selection criterion is the Akaike information criterion (AIC) or its small-sample analog, the corrected Akaike information criterion AIC$_c$ [29]. It evaluates the model response residuals as well as the model complexity in terms of the dimension of the model parameter vector:

$$\mathrm{AIC} = -2\mathcal{L}(\hat{\theta}_{\mathrm{MLE}}) + 2K, \tag{19}$$

with $\mathcal{L}$ the log likelihood function, $\hat{\theta}_{\mathrm{MLE}}$ the maximum likelihood estimated model parameters [51], and K the number of estimated parameters. The corrected AIC is

$$\mathrm{AIC}_c = \mathrm{AIC} + \frac{2K(K+1)}{N-K-1}, \tag{20}$$

with N the number of data points, where its use is recommended for N/K ratios below 40 [29]. From Eq. (19), it is obvious that the AIC penalizes the goodness-of-fit measure, i.e., the log likelihood, through the number of estimated parameters K. The AIC is a relative measure and therefore can only be interpreted within an isolated study, with lower values being favorable [29]. In the developed strategy, the AIC$_c$ serves as a heuristic and supports the model selection procedure. Moreover, two common measures to judge the goodness-of-fit are reported. That is, the root mean square error

$$RMSE = \sqrt{\sum_i^N \frac{\left(c_i^{\mathrm{exp}} - c_i^{\mathrm{sim}}\right)^2}{N}}, \tag{21}$$

and the averaged relative deviation of artemisinin

$$\Delta_{\mathrm{ART}} = \frac{1}{N_{\mathrm{ART}}} \sum_i^{N_{\mathrm{ART}}} \frac{\left|([\mathrm{ART}]_i^{\mathrm{exp}} - [\mathrm{ART}]_i^{\mathrm{sim}})\right|}{[\mathrm{ART}]_i^{\mathrm{sim}}} \tag{22}$$

between simulated (sim) and experimental (exp) data. For the latter, points smaller than 0.01 mol L^{-1} are excluded as they would distort the measure.

The likelihood used for parameter estimation and the AIC (Eq. (19)) was based on independent and identically distributed Gaussian additive errors with zero mean [51]. Thus, the number of estimated parameters K is the sum of the number of model parameters q from Eq. (18) plus one for the measurement variance, $K = q + 1$. The parameters were estimated using Julia [52] from 400 continuous and 201 batch data points (N=601).

The building of different model candidates was based on the available information about the mechanisms in literature as discussed in Section 2, expert domain knowledge and experiences gained from the conducted experiments (Section 4.1). We began with a simple initial base model and progressively explored more complex model candidates to determine whether the increased complexity provided a better description of the observed experimental behavior. The process of identifying the kinetic model from the set of the postulated model candidates by applying the outlined identification strategy is demonstrated in the upcoming results sections starting with a data analysis of the experimental data.

4. Results

The results chapter, first, addresses the outcomes of the continuous and batch experiments, along with the modeling challenges and insights they provide. Next, the kinetic network is introduced, starting with a simple base network and potential extensions to better reproduce the reaction system's behavior. The corresponding kinetic models are subsequently fitted to the experimental data, highlighting which characteristics are captured well and which poorly. Finally, the best-fit model candidate is presented, followed by a detailed discussion of its validation procedure and limitations.

4.1. Experimental data analysis

The investigation of the reaction kinetics of the acid-catalyzed reaction sequence is based on two separate sets of experimental data. One data set was collected in the continuous setup in the presence

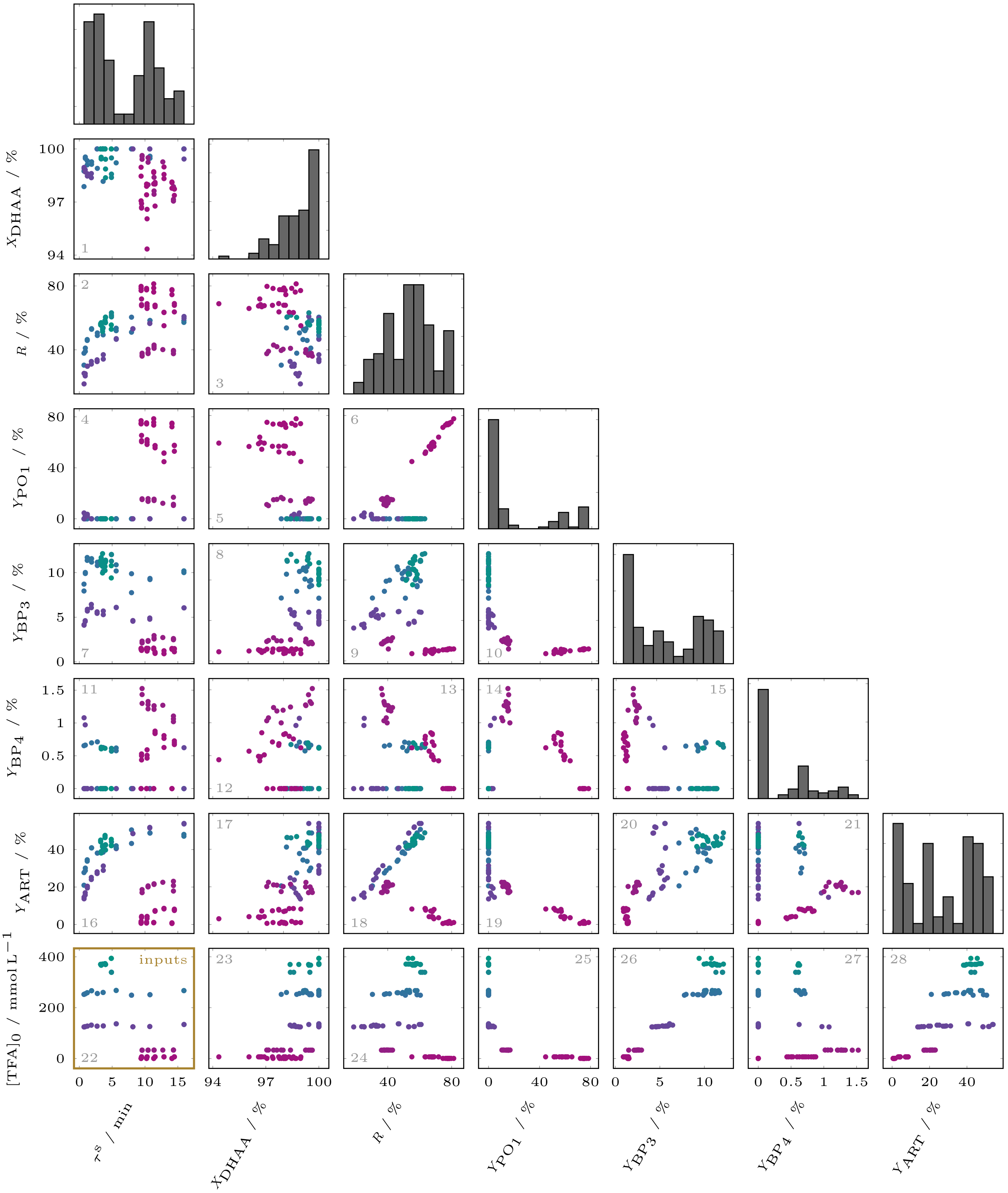


Fig. 6. Scatter matrix and histograms of the continuous setup data with variation in superficial residence time τ^s and initial TFA concentration $[TFA]_0$ (plot 22) for the acid-catalyzed sequence step of the ART synthesis. Please note that the photoreactor was connected upstream of the synthesis reactor to form the hydroperoxide PO_1. The data points are colored according to their acid concentration, see bottom row. The yields Y are based on the converted DHAA concentration. The recovery R considers all recovered species.

of oxygen while operating in steady state mode. The second data set results from experiments in a batch setup for time-resolved sampling that were run in the absence of oxygen. Besides the availability of oxygen, the major differences between the two data sets are the amount of dosed TFA ($[TFA]_0$) and the residence time: we use the actual residence time t in the batch setup and the superficial residence time τ^s, Eq. (1), in the continuous setup. The TFA concentration for the continuous experiments was one magnitude larger than for the batch

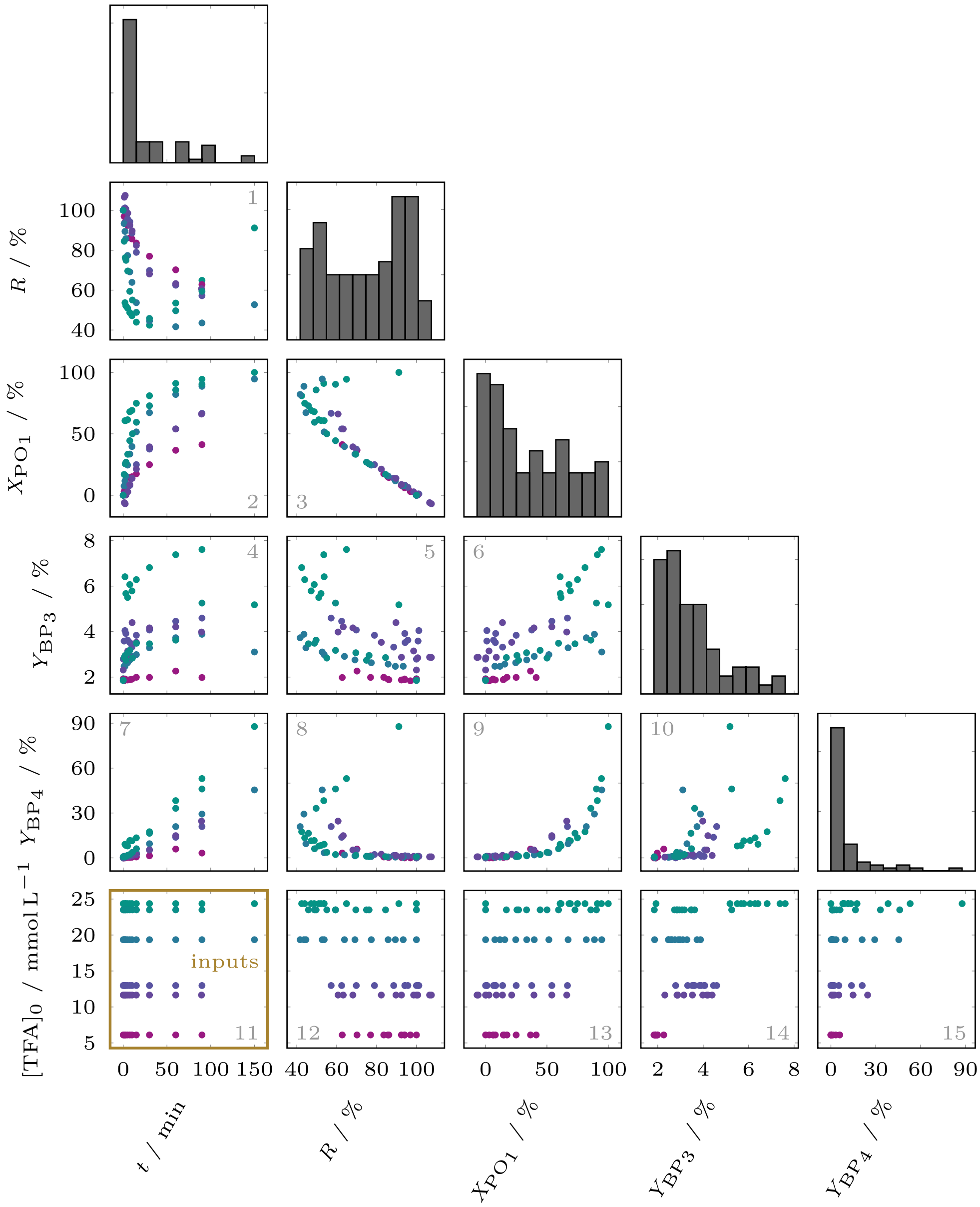


Fig. 7. Scatter matrix and histograms of the batch setup data with variation in initial TFA concentration $[TFA]_0$ and sample time t for the acid-catalyzed sequence step of the ART synthesis. The data points are colored according to their acid concentration, see bottom row. Note two subtle differences between this Figure and Fig. 6: (i) the units in the initial TFA concentration, $\mathrm{mmol\,L^{-1}}$ in this Figure versus $\mathrm{mol\,L^{-1}}$ in Fig. 6; and (ii) the time variables, i.e., the actual sample time t in this Figure versus the superficial residence time τ^s in Fig. 6.

experiments thereby enabling shorter residence times. Figs. 6 and 7 provide scatterplot matrices showing the key quantities from Section 3.2 for the participating species and for the continuous and the batch data, respectively. The scatterplots are numbered in gray inside each plot to guide the discussion. The shown data are colored according to the TFA concentration; see the bottom rows of both Figures.

First, the continuous data visualized in Fig. 6 are analyzed. For the various initial TFA concentrations and flow rates, i.e., residence times, (plot 22) high DHAA conversions of larger than 94 % are achieved for all the continuous experiments (plot 1). ART yields reach 54 % and 47 % within 16 min theoretical residence time for 0.125 M and 0.25 M TFA, respectively, plot 16. The significant difference in ART yields emphasizes the effect of the catalyst concentration on the formation of potentially unknown byproducts. Recoveries on the other hand only reach around 60 % when production of ART has reached a plateau, see green and blue marks in plot 2 and 16. Consequently, more than 40% of the converted DHAA are not identified during the applied measurement procedure and are therefore not known. Moreover, since the recovery significantly drops below 40% for low residence times, plot 2, this suggests at least one non-detected intermediate between the reaction path from PO_1 to ART in Fig. 2 next to the existence of non-quantified byproducts. From the ^{1}H−NMR spectra, none of the other previously reported species could be inferred (details in [28]). This finding is mainly due to the fact that the quantification of many different species with structural similarities is limited with the used analytical methods [28]. Please note that the formation of the measured arteannuin H, one of the major byproducts in the semi-synthesis [28], is not considered in detail, as it is a loss product of a hydroperoxide during the photooxygenation and therefore not relevant to the conversion of PO_1 to ART. Generally speaking, BP_3 is formed in one magnitude greater than BP_4 in the continuous experiments, plots 7 and 11. More specifically, BP_4 only appears in trace quantities, causing the BP_4 data to suffer from a high signal-to-noise ratio. Correlation patterns for BP_4 are therefore difficult to detect in the plots, with the exception of the

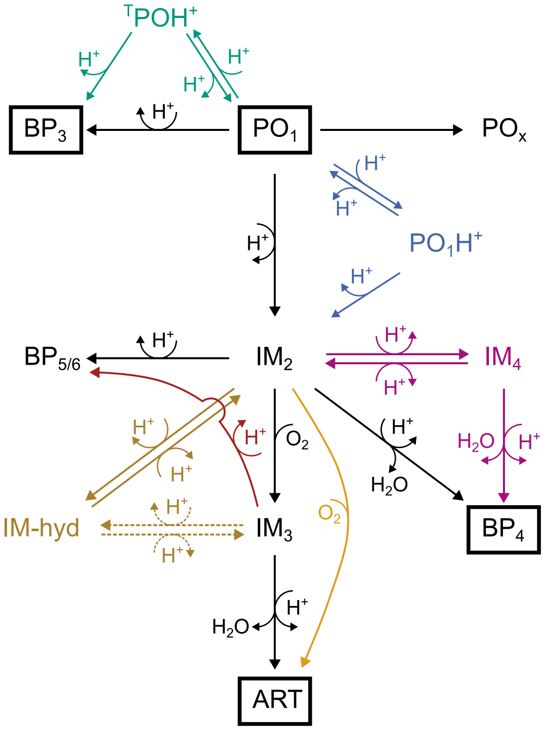


Fig. 8. Simplified superstructure for the acid-catalyzed sequence producing artemisinin from the hydroperoxide PO_1. The base network, i.e., model candidate 1 is drawn in black. Alternative routes are colored, and IM-hyd can either be formed from IM_2 or IM_3 (dashed route).

series performed with the lowest TFA concentration, which suggests a linear correlation between the BP_4 and the ART yield, plot 21. In the case of BP_3, a clear linear correlation with the TFA concentration is observed, plot 26. A correlation with TFA can also be seen for ART, plot 28, and between BP_3 and ART, plot 20. From plot 7 it might be concluded that the formation of BP_3 happens very fast, as the BP_3 yields show a very sharp rise at low residence times reaching quickly a plateau. This indicates an instantaneous appearance of BP_3, i.e., the reaction path towards BP3 must exhibit very fast kinetics.

The complementary batch results were run six times at four different TFA concentrations and for times up to 150 min (plot 11 in Fig. 7).

A comparison between the batch and the continuous results at similar acid concentrations and initial conditions yield similar concentrations for PO_1 and BP_3 suggesting that results from the two setups are comparable (see Appendix B). At the same time, measurement results within batch experiments obtained at equal TFA amounts do show discrepancies, that make drawing conclusions more difficult, see for example the BP_3 yield in plot 4 for the two similar TFA concentrations around 25 $\mathrm{mmol\,L^{-1}}$. These discrepancies might be attributed to deviations in the addition of the acid and an initial temperature rise due to the exothermic nature of the reaction [28]. The complete disappearance of PO_1 takes more than 90 min at the considered TFA concentrations, plot 2. Similar to the continuous data, recoveries can significantly drop with the conversion of PO_1 below 60%, plot 1 and plot 3, and start rising again as BP_4 is increasingly formed when time advances, plot 7. This observation underscores the presence of unidentified intermediates in the pathway leading to ART. In contrast to the continuous case, the main byproduct is BP_4, whose concentrations are an order of magnitude larger than those of BP_3. The buildup of BP_4 is not finished within the considered run time of up to 150 min, plot 7. In the same plot, a strong linear dependence of BP_4 on the residence time is observed. A dependence of PO_1 and BP_4 on TFA can be observed as well, plot 13 and 15, respectively. On the other hand and in contrast to the continuous data, a dependence of BP_3 on the acid concentration cannot be concluded with certainty, plot 14. Interestingly, a distinct non-linear dependence between PO_1 and BP_4 yields can be observed that might be independent on the acid concentration, plot 9.

Table 1
Goodness of fit for the continuous and batch process models describing the acid-catalyzed sequence with the base model, i.e., model candidate 1; details see Table D.6.

Model-data fit			
Symbol	Unit	Value	Description
$N_{\mathrm{conti/batch}}$	1	400/201	Number of data points
$RMSE$	mol/L	0.029	Root mean square error, Eq. (21)
$RMSE_{\mathrm{conti}}$	mol/L	0.027	
$RMSE_{\mathrm{batch}}$	mol/L	0.032	
$\varDelta_{\mathrm{ART}}$	%	77.20	Averaged relative deviation of ART, Eq. (22)
Estimated parameters			
Symbol*	Unit	Value	
k_{PO_x}	$\mathrm{min^{-1}}$	6.14×10^{-4}	
$k_{\mathrm{PO_1 \cdot BP_3}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	0.349	
$k_{\mathrm{PO_1 \cdot IM_2}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	3.17	
$k_{\mathrm{IM_2 \cdot BP_4}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	0.441	
$k_{\mathrm{IM_2 \cdot BP_5}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	0.0	
$k_{\mathrm{IM_2 \cdot IM_3}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	205.0	
$k_{\mathrm{IM_3 \cdot ART}}$	$\mathrm{L\,mol^{-1}\,min^{-1}}$	0.598	

*: The kinetic constant $k_{B \cdot C}$ corresponds to the reaction that forms species C from species B.

Summing up, the key findings from the experimental data analysis are:

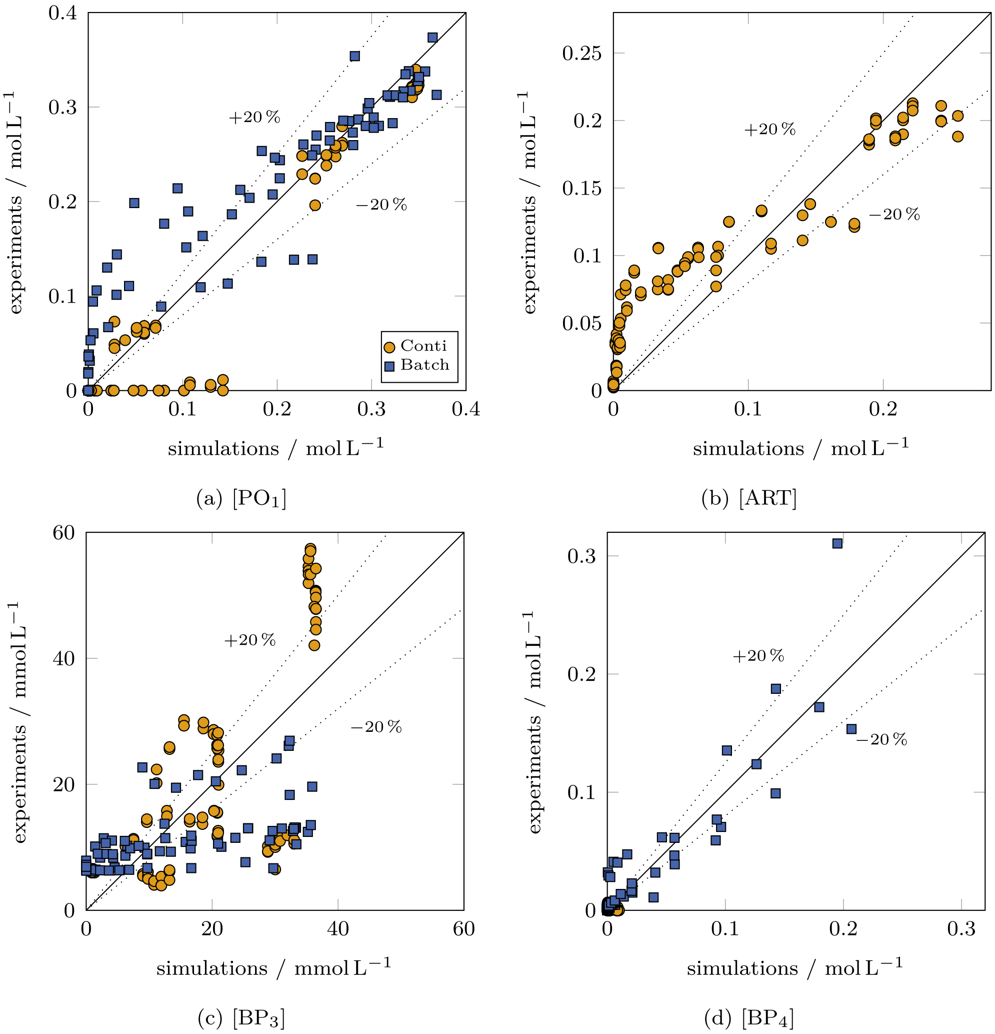
- The batch and the continuous experiments yield similar concentrations for PO_1 and BP_3 at similar TFA and initial concentrations.
- Residence times for the batch data are in the range of hours, and for the continuous data in the range of minutes, mainly caused by a difference of one magnitude in the TFA concentration.
- For continuous experiments reaching a steady state, around 40% of the added DHAA cannot be recovered. This is attributed to non-identified byproducts and to at least one non-identified intermediate on the reaction path from PO_1 to ART.
- The main identified byproduct is BP_3 and BP_4 in the continuous and in the batch data, respectively. In both cases, the more abundant species concentration is one magnitude larger than the concentration of the other byproduct.
- The TFA concentration shows a large influence on the PO_1, BP_3 and ART concentrations in the continuous data, and less pronounced on the PO_1 and BP_4 concentrations in the batch data.

In the following sections, the kinetic model identification strategy summarized in Fig. 5 is applied to the analyzed data sets leading to a kinetic model for the acid-catalyzed sequence of the ART partial synthesis.

4.2. Postulated model candidates

The starting point of the reaction network development was the reaction mechanism of the conversion of PO_1 to artemisinin, introduced in Section 2. Fig. 8 depicts the developed base network for the acid-catalyzed sequence together with the extensions added later. The base model is drawn in black and alternative routes are colored. All tested networks are based on the following assumptions and simplifications:

- The formation of all end products, e.g., ART, BP_3, BP_4 and BP_5 (Fig. 8), is irreversible. They are chemically stable under the conditions applied and do not decompose.
- During the reaction to BP_3, hydrogen peroxide is formed. It does not influence the reaction kinetics and is therefore not considered.



(a) [PO_1]
(b) [ART]
(c) [BP_3]
(d) [BP_4]

Fig. 9. Match of experimental concentrations of (by) products with simulated concentrations based on the parametrized base process model (model candidate 1): (a) PO_1, (b) ART, (c) BP_3, (d) BP_4. The orange circles correspond to the continuous and the blue squares to the batch experiments, respectively. The data of the continuous experiments was sampled at the reactor outlet, whereas the data of the batch experiments was collected in a time-resolved manner. The dashed lines mark 20 % deviations.

- Each species is protonated only once. If not specified otherwise, the acid catalyst reaction order is unity, that is the rate of reaction *r* is proportional to the proton concentration.
- Only TFA contributes to the acidity of the reaction solution, i.e., the carboxyl group of PO_1 does not interfere.

PO_x is a byproduct species we introduced in the previous paper on the photooxygenation kinetics [23]. It includes all formed hydroperoxides formed in the photooxygenation of DHAA apart from PO_1, and their corresponding reaction products after conversion in acidic conditions, e.g. arteannuin H. In the base network, PO_1 undergoes irreversible acid-catalyzed conversion to either the byproduct BP_3 or the intermediate IM_2, with protonation as the rate-limiting step in both pathways. The formation of several intermediate species is simplified to a single lumped species, IM_2, which represents the central enol intermediate and the precursor for both artemisinin and BP_4. The formation of BP_4 is assumed to be governed by a single acid-catalyzed, rate-determining step involving water elimination. A second byproduct, BP_5, is introduced to account for uncharacterized byproducts and low recovery observed in the experimental data. IM_2 irreversibly reacts with oxygen to form IM_3, which is subsequently converted to artemisinin in an acid-catalyzed step that eliminates water. This simplified mechanism assumes oxygen addition as the rate-determining step, with irreversible water abstraction. This base model candidate contains 4 byproducts and 7 reaction steps and assumes a reaction order of $m=1$ regarding acid concentration for the acid-catalyzed steps.

In later model candidates, additional reaction pathways were incorporated to account for reversible steps and side reactions. A reversible protonation step was added to the cleavage of PO_1, where protonation at either the internal or external oxygen of the hydroperoxyl group leads to a catalyst-binding intermediate and a subsequent forward reaction releasing the catalyst [53]. This protonation pathway substitutes the direct conversion of PO_1 to IM_2 and BP_3. Further, an internal equilibrium step was introduced, where IM_2 reversibly forms IM_4, which then irreversibly transforms into BP_5. IM_4 is interpreted as the seco-cadinane species, a major byproduct identified in prior studies. Finally, water effects were considered by introducing hydration of IM_2 and IM_3, forming IM-hyd. As the reaction progresses, the increased water concentration shifts the equilibrium towards IM-hyd, slowing both BP_4 and artemisinin formation. The details of the tested model candidates are given in Table D.6.

4.3. Performance of the base model

In the following, the results of the base model candidate (Fig. 8 in black) are shown first before the identification path to the final kinetic model considering the alternative routes is discussed.

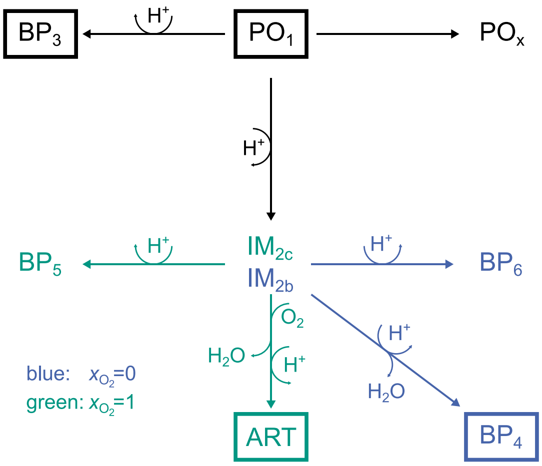


Fig. 10. Selected reaction network for the acid-catalyzed sequence producing artemisinin from the hydroperoxide PO_1. The network consists of two sub-networks. The green sub-network is active when the gas phase consists of oxygen (x_{O_2}=1, continuous experiments). The blue sub-network is active when no oxygen is present (x_{O_2}=0, batch experiments).

Table 2
Goodness of fit and estimated parameter values and spreads for the developed process model and selected and refined model candidate 16 describing the synthesis reaction network.

Model-data fit

Symbol	Unit	Value	Description
$N_{\text{conti/batch}}$	1	400/201	Number of data points
$RMSE$	mol/L	0.016	Root mean square error, Eq. (21)
$RMSE_{\text{conti}}$	mol/L	0.014	
$RMSE_{\text{batch}}$	mol^2/L^2	0.021	
Δ_{ART}	%	13.97	Averaged relative deviation of ART, Eq. (22)

Estimated parameters

Symbol*	Unit	Value	CI_{95}^{-} †	CI_{95}^{+} †
k_{PO_x}	1/min	0.0105	0.0096	0.0113
m_{IM_2}	–	2.646	2.629	2.664
$k_{PO_1 \cdot BP_3}$	$(\text{L/mol})^{m_{IM_2}+1}/\text{min}$	424.2	373.0	477.1
$k_{PO_1 \cdot IM_2}$	$(\text{L/mol})^{m_{IM_2}}/\text{min}$	880.3	833.2	931.6
$k_{IM_{2b} \cdot BP_6}$	$\text{L}\,\text{mol}^{-1}\,\text{min}^{-1}$	0.0	–	–
$k_{IM_{2c} \cdot BP_5}$	$\text{L}\,\text{mol}^{-1}\,\text{min}^{-1}$	0.682	0.614	0.755
$k_{IM_{2b} \cdot BP_4}$	$\text{L}\,\text{mol}^{-1}\,\text{min}^{-1}$	0.887	0.770	1.026
$k_{IM_{2c} \cdot ART}$	$(\text{L/mol})^2/\text{min}$	30.07	28.07	32.24

*: The kinetic constant $k_{B \cdot C}$ corresponds to the reaction that forms species C from species B.

†: $[CI_{95}^{-}, CI_{95}^{+}]$: 95% confidence interval based on profile likelihood [23].

The results of fitting the base model to the experimental data are shown in Fig. 9 and Table 1. The averaged relative deviation of ART almost reaches 78 %, suggesting that the base model is not representing fully the reaction network underlying the experiments. From the estimated kinetic parameter for the reaction of intermediate IM_2 to BP_5, $k_{IM_2 \cdot BP_5}=0\,\text{L}\,\text{mol}^{-1}\,\text{min}^{-1}$, it follows that the byproduct reaction yielding BP_5 is not present. The high value for $k_{IM_2 \cdot IM_3}=205\,\text{L}\,\text{mol}^{-1}\,\text{min}^{-1}$ suggests that IM_2 is converted almost instantly to IM_3 in the presence of oxygen. In the PO_1 parity plot for the continuous data, Fig. 9(a), the model predicts for some measurements non-vanishing amounts of PO_1, although the data suggest near-zero concentrations; see the data points on or very close to the abscissa. At the same time, ART concentrations for relatively low ART values are predicted too low, Fig. 9(b). For the batch data, an opposite trend is visible: The base model does predict zero or trace quantities of PO_1, but the experimental data shows significantly higher concentrations. Thus, the PO_1 depletion is predicted considerably too fast. The data points for ART in Fig. 9(b) form a curvy pattern that suggests a systematic gap in the base model structure. Furthermore, the BP_3 data, Fig. 9(c), show very poor predictive capabilities of the base model over almost the entire concentration range, which points to missing characteristics of the simulated chemical reaction network. Hence, the base model is not able to predict the measured species within the considered concentrations, making extensions and variants of the base model necessary.

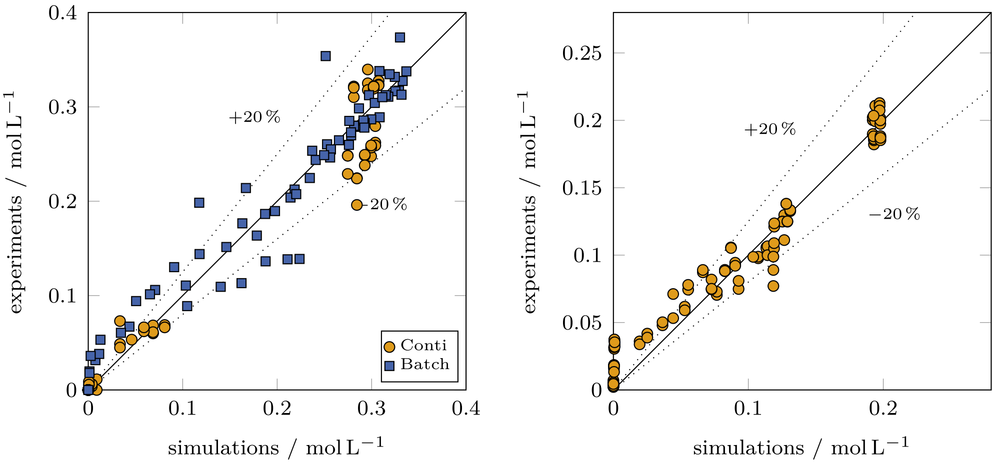
4.4. Route to identify and parameterize a final kinetic model

Different combinations of reactions of the full reaction network shown in Fig. 8 have been analyzed to identify the most likely candidate. Information about the different model candidates built upon these modifications can be found in Table D.6 in Appendix. Likewise, the root mean square errors and the corresponding AIC_c values resulting from the parameter identification are given in Table D.7 in Appendix. In the following, the key results of the analysis that led to the final kinetic model are summarized. The model candidates are thereby referenced by their unique number as specified in Table D.6.

Changing the dissociation behavior of TFA from the default full dissociation to partial dissociation (see equations in Table D.5) does not improve the model-data fit (model candidates 1, 2, 3, and 16, 17, 18 in Table D.6). Model candidates with different schemes for protonation, dissociation equilibrium, hydration and byproducts that consider proton reaction orders of unity (candidates 4 to 12) do not lead to a notable improvement of the model-data fit in terms of the root mean square error ($RMSE$). The largest improvement is achieved by assuming a protonation step for PO_1 reacting to IM_2 (blue modification in Fig. 8; model candidates 4, 5, 7, and 10 to 12) emphasizing the important influence of TFA on the reaction kinetics. Additionally integrating hydration of IM_2 before it can react to IM_3 (brown modification in Fig. 8; model candidates 10 and 12) yields the best performing model within this group having an $RMSE$ as low as $0.018\,\text{mol}\,\text{L}^{-1}$. On the other hand, adding a further intermediate IM_4 for the production of BP_4 from IM_2 (purple modification in Fig. 8; model candidate 12) shows a negative effect on the predictive capability of the model, so this reaction is unlikely. Clearly, the key aspects beneficial for the improved model-data fit in model candidate 10 are the variability in the acid concentration over the course of the reaction, and the possibility of a resting state by the hydration extension. However, the usage of less TFA in the batch experiments to reduce the reaction rate resulted in almost complete temporary TFA depletions in the batch experiments for model candidate 10. This was considered unrealistic, in particular when compared with the TFA behavior in the continuous experiments. Consequently, the alteration of the TFA reaction orders for the depletion of PO_1 towards both BP_3 and IM_2 were targeted (model candidates 13 to 30). For all model candidates within this category, significant improvements are achieved, where most of them reach an $RMSE$ of $0.019\,\text{mol}\,\text{L}^{-1}$ or below. The apparent reaction orders of TFA lie between 2 and 3 for m_{IM_2} and 3 to 4 for m_{BP_3}. Thus, the TFA reaction orders considerably exceed 1. Such high values might be a consequence of a sequence of rate-limiting acid-catalyzed steps, i.e., multiple protonations lumped together; compare with the mechanism in Fig. 2. Alternatively, the high reaction orders compensate physical or chemical phenomena, that are not covered by the structures of the model candidates.

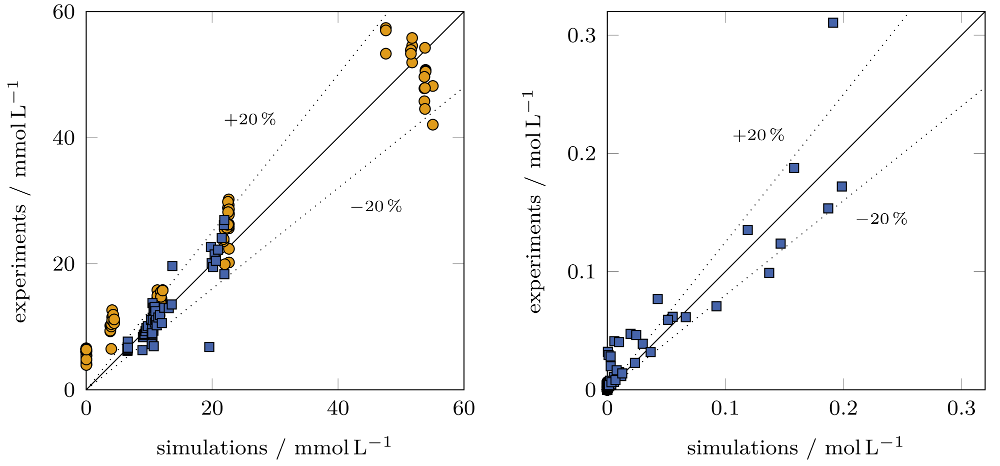
In many model candidates, the formation of IM_3 was estimated to be very fast. That is, IM_3 is accumulated as the main intermediate and IM_2 is only formed in trace quantities in the presence of oxygen leaving no reactant for the byproduct formation starting from IM_2. Therefore, the byproduct formation starting from IM_3 obtains better model-data fits (see model candidate pairs 15/16 and 25/26). This suggests the existence of an additional byproduct on the reaction path from IM_2 to ART apart from BP_4.

The consideration of further structural extensions and variants, other than the modifications shown in Fig. 8, did not improve the



(a) [PO_1]

(b) [ART]



(c) [BP_3]

(d) [BP_4]

Fig. 11. Match of experimental concentrations of (by) products with simulated concentrations: (a) PO_1, (b) ART, (c) BP_3, (d) BP_4. Simulations are based on the final kinetic model, Eqs. (24), combined with the reactor model (Section 3.3 and Appendix C in the Appendix). The dashed lines mark 20 % deviations.

model-data fit and are therefore not discussed here [48]. From the postulated kinetic model candidates, the most promising one supported by the magnitudes of the AIC_c was chosen. From Table D.7 in the Appendix, a small choice of model candidates reaches an AIC_c of almost −5000, with very subtle differences among the candidates (see the models with $AIC_c < -4890$). It is known that AIC/AIC_c tends to overfitting, i.e., prefers model candidates with a greater number of model parameters [54]. Therefore, model candidate 16 was chosen that has both a small number of parameters, namely 7, a good model-data fit and a low AIC_c. Here, PO_1 converts to BP_3 and IM_2 directly (not via the blue or green routes in Fig. 8) with non-unity TFA reaction orders, where the reaction order to BP_3 is higher by 1 than that for IM_2. IM_2 is either oxygenated to IM_3, or converted to the byproduct BP_4, i.e., hydration of IM_3 (brown variant) and IM_4 (violet variant) are discarded. IM_3 can react to a further byproduct BP_5 (red-brown route; reaction from IM_2 to BP_5 is discarded) or to the final product ART.

4.5. Further refinements of the final kinetic model

Model candidate 16 underwent two further model refinements: The first refinement was inferred from interpreting the model-data fit while considering the partial synthesis mechanism. The second refinement was established on the basis of identifiability problems of model parameters.

Considering the continuous and the batch data of the BP_4 species, Figs. 6 and 7, BP_4 is almost exclusively formed in the batch experiments. On the other hand, the model candidates generally overestimated the amount of BP_4 formed in the continuous experiments and underestimated the BP_4 formation in the batch experiments. An important difference in the prevalence of reactions between the continuous and batch experiments is related to the impact of IM_2 to produce BP_4. In the batch experiments, in contrast to the continuous case, the reaction path from IM_2 towards ART is blocked as the batch reactor is run in oxygen-free atmosphere. Albeit, BP_4 can and is predicted to form in the continuous experiments, even in situations where the model candidates showed a very short occurrence of IM_2. Hence, the results from the model selection revealed that the differing formation rates of BP_4 from IM_2 in the continuous and batch experiments could not be sufficiently explained by the identified joint model.

$k_{IM_2 \cdot IM_3}$ is the only estimated parameter that reached its preset upper limit, i.e., 5000 L mol^{-1} min^{-1}. This is equal to a reaction from IM_2 to IM_3, which is solely limited by oxygen mass transfer. The model-data fit could be further improved by increasing the mass transfer coefficient to physically unrealistic numbers resulting in an instant conversion of IM_2 to IM_3 in the presence of oxygen. Alternatively, two different intermediates for IM_2 in the continuous and batch case, namely IM_{2c} and IM_{2b}, were introduced, that improved the model-data fit and eased the numerical simulations as the rapid formation of IM_3 in the continuous experiments was eliminated. The introduction of two different intermediates for the same physical system is the trade-off one must accept for approximating the real system behavior by a simplified reaction scheme that is to be identified from two limited data sets collected under different conditions. A physical interpretation is the accumulation of two different intermediates in the batch and continuous experiments. Note however that the differences in the model-data fit of model candidate 16 between its original and its refined version are subtle. Hence, the original version can serve as a generic model in situations where the experimental conditions fall in between the studied batch and continuous case. That is, for example, if the molar oxygen fraction in the gas phase is between 0 (batch case) and 1 (continuous case).

The second refinement compensates for a detected parameter identifiability issue [51] that all of the model candidates shared. An analysis revealed that a non-identifiability is linked to the kinetic constant $k_{\mathrm{IM_{2c}\cdot IM_3}}$ [48] that therefore cannot be estimated. Thus, only one intermediate is accumulated in non-trace quantities. This leads to the omission of $\mathrm{IM_3}$ and shortens therefore the species number of the pathway to ART by one. The omission of $\mathrm{IM_3}$ also results in smoother liquid oxygen kinetics, where in the presence of $\mathrm{IM_3}$ a rapid depletion of liquid oxygen is predicted. It occurs as a result of very high magnitudes for $k_{\mathrm{IM_{2c}\cdot IM_3}}$, as the oxygenation was assumed to happen from $\mathrm{IM_2}$ to $\mathrm{IM_3}$, Fig. 8.

The final reaction network is depicted in Fig. 10. Key characteristic of the network is that different sub-networks are active depending on the gas phase composition. If the gas phase contains oxygen ($x_{\mathrm{O_2}}$=1, green sub-network in Fig. 10), the accumulating intermediate is $\mathrm{IM_{2c}}$, and ART and byproduct $\mathrm{BP_5}$ are formed, but no $\mathrm{BP_4}$. On the other hand, if oxygen is absent ($x_{\mathrm{O_2}}$=0, blue sub-network in Fig. 10), the accumulating intermediate is $\mathrm{IM_{2b}}$, and $\mathrm{BP_4}$ and $\mathrm{BP_6}$ are the only products formed. The kinetic model corresponding to the reaction network for the acid-catalyzed sequence in Fig. 10 using mass action kinetics, Eq. (5), reads

$$\mathrm{PO_1} \xrightarrow{} \mathrm{PO}_x, \tag{23a}$$

$$\mathrm{PO_1} \xrightarrow{\mathrm{TFA}} \mathrm{BP_3}, \tag{23b}$$

$$\mathrm{PO_1} \xrightarrow{\mathrm{TFA}} \mathrm{IM_{2c/2b}}, \tag{23c}$$

$$\mathrm{IM_{2c}} \xrightarrow{x_{\mathrm{O_2}} = 1,\,\mathrm{TFA}} \mathrm{BP_5}, \tag{23d}$$

$$\mathrm{IM_{2b}} \xrightarrow{x_{\mathrm{O_2}} = 0,\,\mathrm{TFA}} \mathrm{BP_6}, \tag{23e}$$

$$\mathrm{IM_{2b}} \xrightarrow{x_{\mathrm{O_2}} = 0,\,\mathrm{TFA}} \mathrm{BP_4} + \mathrm{H_2O}, \tag{23f}$$

$$\mathrm{IM_{2c}} + {}^3\mathrm{O_2} \xrightarrow{\mathrm{TFA}} \mathrm{ART} + \mathrm{H_2O}, \tag{23g}$$

with rates of reactions r_i:

$$r_{\mathrm{PO_1\cdot PO_x}} = k_{\mathrm{PO_1\cdot PO_x}}[\mathrm{PO_1}] \tag{24a}$$

$$r_{\mathrm{PO_1\cdot BP_3}} = k_{\mathrm{PO_1\cdot BP_3}}[\mathrm{TFA}]^{m_{\mathrm{TFA}}+1}[\mathrm{PO_1}] \tag{24b}$$

$$r_{\mathrm{PO_1\cdot IM_2}} = k_{\mathrm{PO_1\cdot IM_2}}[\mathrm{TFA}]^{m_{\mathrm{TFA}}}[\mathrm{PO_1}] \tag{24c}$$

$$r_{\mathrm{IM_{2c}\cdot BP_5}} = k_{\mathrm{IM_{2c}\cdot BP_5}}[\mathrm{TFA}][\mathrm{IM_{2c}}] \tag{24d}$$

$$r_{\mathrm{IM_{2b}\cdot BP_5}} = k_{\mathrm{IM_{2b}\cdot BP_5}}[\mathrm{TFA}][\mathrm{IM_{2b}}] \tag{24e}$$

$$r_{\mathrm{IM_{2b}\cdot BP_4}} = k_{\mathrm{IM_{2b}\cdot BP_4}}[\mathrm{TFA}][\mathrm{IM_{2b}}] \tag{24f}$$

$$r_{\mathrm{IM_{2c}\cdot ART}} = k_{\mathrm{IM_{2c}\cdot ART}}[\mathrm{TFA}][\mathrm{IM_{2c}}][\mathrm{O_{2,l}}] \tag{24g}$$

Note that in the presence of oxygen ($x_{\mathrm{O_2}}$=1, continuous case) $\mathrm{IM_{2b}}$ and in the absence of oxygen ($x_{\mathrm{O_2}}$=0, batch case) $\mathrm{IM_{2c}}$ is inactive. Both intermediates are formed with the same rate, where the reaction order of TFA, m_{TFA}, is set as parameter to be estimated. In the formation of $\mathrm{BP_3}$, the reaction order of TFA is set 1 larger than that of $\mathrm{IM_2}$ formation.

4.6. *Performance of the final model*

A quantitative summary of the parameter estimation results of the final model is given in Table 2. A drastic improvement of the model-data fit is observed compared to the results of the base model candidate, Table 1, where the *RMSE* decreased from $0.029\,\mathrm{mol\,L^{-1}}$ to $0.016\,\mathrm{mol\,L^{-1}}$. The continuous data is slightly better predicted, which could be observed for all model candidates (see Table D.6 in the Appendix). The averaged relative deviation of ART is 13.97%, where points smaller than simulated values of $0.01\,\mathrm{mol\,L^{-1}}$ are not considered because they would distort the calculation of the relative measure that is based on the simulated ART concentration. The high value estimated for the apparent reaction order of TFA of 2.46 indicates that the depletion of $\mathrm{PO_1}$ is a sequence of rate-limiting acid-catalyzed steps, i.e., multiple protonations in the mechanisms in Fig. 2 are lumped together. Alternative explanations for the high apparent reaction orders include non-ideal mixing effects, pH-dependent speciation of the catalyst, or variations in activity coefficients at high TFA concentrations [55,56]. Note that these effects are implicitly captured in our effective kinetic parameters. Practically, all eight parameters are identifiable as a corresponding analysis showed [48], with a good confidence in all of them as can be seen in Table 2.

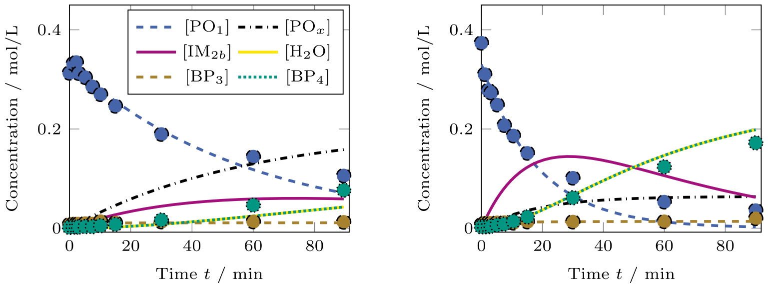
The parity plots of the measured versus simulated data are shown in Fig. 11. In general, the results show a satisfactory agreement between simulated and experimental data considering the complexity of the reaction network and the continuous reactor setup, in particular for the main pathway to ART, Fig. 11(b).

However, there are deviations between model and experimental data for both the continuous and the batch data set. The model could not fully reproduce the $\mathrm{PO_1}$ conversion in the batch data and likewise the build-up of $\mathrm{BP_4}$, especially at concentrations of $\mathrm{PO_1}$ below $0.25\,\mathrm{mol\,L^{-1}}$ (Fig. 11(a)). For the continuous data, the model predicts the formation of ART and $\mathrm{BP_3}$ too slow at low TFA concentrations, as seen in the ART and $\mathrm{BP_3}$ points lying on the ordinate in Figs. 11(b) and 11(c). This behavior results from the slow depletion of $\mathrm{PO_1}$ at low TFA concentrations. The parity plots also show pronounced scattering of experimental data at constant values of the simulated value, e.g. the arrangement of parity points in a vertical manner at high ART and $\mathrm{BP_3}$ concentrations. This is for a large part attributed to the variability of the measurement data obtained in the batch and continuous setup as previously identified in the experimental data analysis in Section 4.1.

4.7. *Generalizability of the identified model*

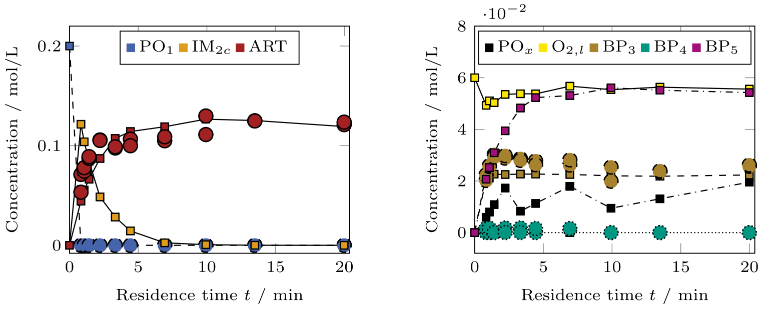
The developed model incorporates several simplifications that are specifically applicable to the system of purified dihydroartemisinic acid in toluene and catalyzed by trifluoroacetic acid. The model is valid only under specific conditions: at −20 °C for the photooxygenation step and at 20 °C for the acid-catalyzed reaction step. Additionally, the application of the model in the acid-catalyzed reaction step is restricted to scenarios where the gas phase consists of either pure oxygen or pure nitrogen.

The application of the drift flux model and the velocity-dependent linear mass transfer approach are strong simplifications of two-phase reactive flows in capillary tubing [57] potentially compromising the isolation of the underlying kinetics from transport phenomena. A major source of concentration gradients are mass transfer limitations. At the reactor inlets, ample tubing precedes both the photooxygenation (3.2 m) and synthesis (0.4 m) reactors, ensuring dissolved oxygen reaches equilibrium before entering, thereby guaranteeing a well-mixed liquid phase. We demonstrated in our prior publication that mass transfer limitations are negligible for the photoreactor through Hatta number analysis [23]. The Hatta number indicated that the reaction occurs in the kinetic regime rather than being mass transfer-limited. The experimental conditions in this study provide even greater assurance



(a) TFA=0.011 mol L^{-1}.

(b) TFA=0.023 mol L^{-1}.



(c) TFA=0.25 mol L^{-1}; main reaction partners.

(d) TFA=0.25 mol L^{-1}; reaction byproducts and oxygen.

Fig. 12. Time profiles of concentrations for exemplary batch (upper row) and continuous (bottom row) experiments at different conditions. Circles are experimental data points; lines in the upper row and squared markers in the bottom row are simulations using the identified kinetic model, Fig. 10 and Eq. (24). In the continuous case, the data points are given at the synthesis reactor outlet. Hence, the time-resolved graphs originate from different experiments. The corresponding residence times were taken from simulations.

of kinetic control, as the pure oxygen atmosphere (versus nitrogen–oxygen mixtures in our previous work) enhances oxygen transfer and diminishes mass transfer limitation risks. It can be concluded that mass transfer limitations play a minor role under the studied conditions, which we confirmed by another Hatta analysis as outlined in the Appendix E. Despite the absence of mass transfer limitations, the necessity of fixing the mass transfer coefficient to the photooxygenation value due to identifiability issues warrants further investigation.

While a general transferability of the kinetic model is limited, the observed kinetic behavior could neither be explained by mass transport limitations nor by the mechanisms and kinetics published in the literature. Nonetheless, a more detailed and microscopic examination of hydrodynamics and transport phenomena would undoubtedly enhance future advancements in the process model, enabling the inclusion of conditions beyond those explored in this study. The reported kinetic parameters represent effective values for the described reactor configuration. This is in particular evident in the apparent reaction order of the acid catalyst being greater than two. The high value reflects the semi-empirical nature of our simplified process description and mass action law formulation rather than true molecular reaction orders. While mass transfer limitations appear minimal based on our analysis, they cannot be entirely ruled out. Consistent with standard practice in process systems engineering, recalibration of parameters may be required for fundamentally different reactor geometries or operating regimes.

To improve the predictive accuracy of the kinetic model, additional research is required. A thorough identification of all reaction products and intermediates is crucial for developing a more widely applicable kinetic model. This will also facilitate the determination and quantification of the mechanistic and kinetic role of the acid catalyst. Moreover, quantifying water formation during the reaction is another critical area of focus. As the reaction mechanism becomes better understood, the model can be expanded to incorporate temperature dependence, as temperature influences various aspects of the process, including transfer phenomena, flow conditions, and reaction kinetics. A key challenge will be to identify the parameters most sensitive to temperature changes, thereby reducing the number of variables that need to be estimated.

4.8. Model application: Analysis of reaction dynamics

The discussed results in the previous Section enable a model-based analysis of the partial synthesis reaction dynamics with the identified process models and their underlying identified kinetic models. For the operation in batch mode, time profiles of two exemplary experiments are shown for two different TFA concentrations in the top row of Fig. 12. Predictions agree in general very well with the experimental data. A minor exception is observed for the PO_1 depletion, where predictions tend to be too fast with proceeding time. This can be observed in all batch experiments.

For the continuous experiments, results from simulations and experimental data having a TFA concentration of 0.25 mol L^{-1} are given in the bottom row of Fig. 12. Overall, the agreement between simulated and experimental data is good. In particular, PO_1 conversion proceeds very fast according to the data, which is well matched by the model. The variation observed in the calculated formation of PO_x lumped species results from the specific residence time set in the photoreactor for each experimental condition. In all experiments, a minimum residence time was established to ensure complete conversion of DHAA. To achieve longer residence times in the subsequent acid-catalyzed section, the flow rate was reduced, which also led to increased residence times in the photoreactor and, consequently, higher byproduct formation. IM_{2c}, the major intermediate, is very rapidly built up, and slightly less rapidly converted to ART. Despite the high rate of ART formation, the liquid oxygen concentration $O_{2,l}$ does not drop significantly (Fig. 12(d)). Thus, the build-up of ART is not mass-transfer limited. As mentioned in the previous Section, the model slightly underestimates the ART and BP_3 formation in the beginning (Fig. 12(c)).

This data also reveals why two separate intermediate species in the network are required to reproduce both the continuous and the batch experimental data. With the final model, IM_{2b} and IM_{2c} are formed with the same rate from PO_1 accumulating as intermediates before being converted with separate rates to BP_4 and ART, respectively. If BP_4 and ART evolve via the same major intermediate IM_2, the rate of BP_4 formation observed in the batch data imposes significant BP_4 product concentrations in the continuous setup despite the short occurrence of IM_2. That is not covered by the experimental data.

5. Conclusion

Designing and optimizing a process for the artemisinin partial synthesis requires a simple but sufficiently accurate kinetic model. Developing a broadly applicable model for the acid-catalyzed sequence from the intermediate PO_1 to artemisinin is particularly challenging due to the complex reaction mechanism, which involves multiple reversible acid-catalyzed steps, unstable intermediates, oxidation reactions, and is coupled with mass transfer. Only 24 % to 76 % of the reaction products formed could be identified in the experiments. The analysis of the different postulated simplified reaction networks showed that a purely linear dependence of the reaction rates on the present acid concentration is not sufficient to describe the experimental data. The selected final model has high H^+ reaction orders exceeding 2, that underlines the importance of the catalyst and the presence of intermediate steps. Thus, with respect to the conversion of PO_1, utmost caution is required when handling acid concentrations in operational settings. Another major characteristic of the identified kinetic model is that it has different sub-networks for an oxygen-absent (batch case) and oxygen-rich (continuous case) atmosphere. This emphasizes the complexity of the partial synthesis of artemisinin with different pathways dominating depending on the operating conditions. With these main features, the selected model candidate can describe the main process behavior with a mean relative deviation for artemisinin of 14%. Based on the model analysis, efficient mass transfer throughout the entire reaction system is crucial for artemisinin production to minimize byproduct formation in oxygen-deprived reaction zones. Combined with previous work on the step from dihydroartemisinic acid to PO_1 [23], a semi-empirical kinetic model for the ART partial synthesis and a process model for the production of ART via partial synthesis in a milliscale photo-flow reactor are made available with this work.

However, since our study provides effective kinetic parameters estimated with the introduced modeling assumptions, a transfer to fundamentally different reactor types or operating conditions may require reparameterization. In more detail, while the identified kinetic model cannot provide full mechanistic understanding due to its semi-empirical nature and its embedding in a simplified reactor description, it does capture key behavioral tendencies that are valuable for process insight and optimization within the studied parameter range. Moreover, the identified kinetic constants are valid only at the investigated temperature of 20 °C using toluene as solvent and trifluoroacetic acid as catalyst. To increase the predictive capability of the developed model, a complete elucidation of all reaction products and intermediates is pending. Secondly, the determination of the roles of water and acid over the reaction progress is required to detect any effect of water on the reaction mechanism and to validate the assumption of a constant acid concentration and the acid's substantial influence on the reaction. When the underlying phenomena are better understood, the model can be extended to study the effect of temperature and solvent on the reaction kinetics.

CRediT authorship contribution statement

Moritz Schulze: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Susann Triemer:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **René Schenkendorf:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Andreas Seidel-Morgenstern:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Ulrike Krewer:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We gratefully acknowledge the support by the Center of Pharmaceutical Engineering (PVZ), Braunschweig and the Max Planck Society. Moritz Schulze and Susann Triemer are also grateful to the International Max Planck Research School “Advanced Methods in Process and Systems Engineering” (IMPRS ProEng). We also gratefully acknowledge the support of the Otto-von-Guericke University and especially Dr. Liane Hilfert and Sabine Hentschel for performing the NMR analysis of the reaction samples. Open Access funding enabled and organized by Projekt DEAL.

Appendix A. Summary of standard and varied reaction conditions

In Section 3.1 we introduced the experimental setup and the chosen conditions for both the continuous and batch experiments. For clarification and specification of the reaction conditions, Table A.3 summarizes the constant and varied concentrations in the feed solutions, the pumps settings, reactor sizes and the resulting conditions, e.g. the residence time and acid concentration in solution.

Appendix B. Comparison of batch and continuous experiments to study the acid-catalyzed reaction sequence

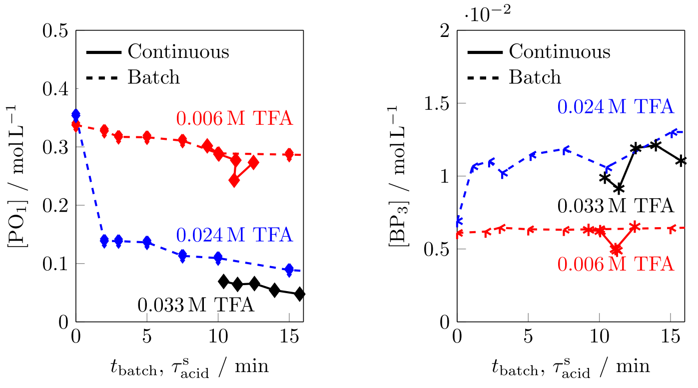
As introduced in Section 3.1, the acid-catalyzed reaction sequence from PO_1 to artemisinin was studied through both continuous experiments in the presence of oxygen and batch experiments under anaerobic conditions. Batch experiments are particularly useful for obtaining multiple kinetic data points in a single run, whereas each continuous experiment in steady-state conditions provides a single data point. Due to the limited residence time in the continuous reactor, batch experiments were employed to investigate lower acid concentrations, which require longer experimental durations.

Fig. B.13 presents a comparison of experimental data from both batch and continuous experiments at similar TFA concentrations. The comparison was made based on the concentrations of PO1 and the byproduct BP3. According to the reaction mechanism (Section 2), the conversion of PO_1 and the formation of BP_3 are unaffected by the presence of oxygen.

The similar results observed after equivalent residence times in the continuous mode and experimental durations in the batch mode indicate that both experimental setups produce comparable outcomes for PO_1 and BP_3, despite the differences in flow conditions.

Table A.3
Reaction conditions of the continuous and batch experiments: Constant and varied properties.

	Continuous setup		Batch setup
	Photoreactor	Acid-catalyzed section	
Temperature	−20°	20 °C	20 °C
Pressure	7 bara		1 bara
Oxygen content in the gas phase	$x_{O_2} = 1$		$x_{O_2} = 0$
Reactant feed composition in toluene	c_{DHAA} = 0.25–0.28 M c_{DCA} = 1.25–1.50 mM		0.2 M PO_1, 2.5 mM DCA
Acidic feed composition in toluene	c_{TFA} = 0.06–4 M		–
Acid concentration in reaction solution	0	c_{TFA} = 0.006–0.4 M	c_{TFA} = 0.006–0.024 M
Reactant-to-acid feed ratio	9:1 (v/v)		–
Gas-to-Liquid ratio	4:1 (v/v)		–
Total flowrate	0.5–2.5 mL min^{-1}		–
Gas flow rate	0.4–2 mL min^{-1}		–
Reactant flow rate	0.09–0.45 mL min^{-1}		–
Acidic flow rate	0.01–0.05 mL min^{-1}		
Total reaction volume	5 mL	1.5–14.5 mL	30 mL
Residence time	2.5–10 min	1–16 min	–
Sampling	–	At the reactor Outlet	After 1, 2, 3, 5, 7.5, 10, 15, 30, 60, 90 min



(a) PO_1

(b) BP_3

Fig. B.13. Comparison of experimental data on the acid-catalyzed reaction sequence obtained in continuous operation in the presence of oxygen and in batch operation in nitrogen atmosphere ($[DHAA]_0$=0.5 M, T_{photo}= − 20 °C, τ^{s}_{photo}=5 min, [DCA]=2.5 mM, x_{O_2}=1, P_{LED}=100 %): (a) Cleavage of PO_1 (b) Formation of BP_3.

Appendix C. Process model for the continuous setup

C.1. Reactions in the photoreactor

The reaction network in the photo reactor from the previously published study [23] is

$$DHAA + {}^1O_2 \xrightarrow{k_{PO_1}} PO_1, \tag{C.1a}$$

$$DHAA + {}^1O_2 \xrightarrow{k_{PO_y}} PO_y, \tag{C.1b}$$

$${}^1O_2 \xrightarrow{1/\tau_\Delta} {}^3O_2, \tag{C.1c}$$

$$PO_1 \xrightarrow{k_{PO_x}} PO_x, \tag{C.1d}$$

$$PO_y \xrightarrow{k_{PO_x}} PO_x, \tag{C.1e}$$

with PO specifying differing (lumped) hydroperoxides and τ_Δ the lifetime of singlet oxygen that is produced by photosensitization.

C.2. Governing equations of the final process model

Combining the reaction network of the photoreactor, Eq. (C.1), with the reaction rates of the identified final model of the acid-catalyzed sequence, Eqs. (24), the governing equations of the process model for the continuous setup are

$$\frac{d[DHAA]}{dz} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_l}\Bigg(-\frac{[O_2]}{k_{l1}[O_2]+k_{l2}}\frac{(\tilde{k}_{PO_1}+\tilde{k}_{PO_y})[DHAA]}{1+(\tilde{k}_{PO_1}+\tilde{k}_{PO_y})[DHAA]}$$
$$L_p(1-\exp[-\kappa c_{DCA} l_{opt}])\Bigg),$$

$$\frac{d[PO_1]}{dz} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_l}\Bigg(\frac{[O_2]}{k_{l1}[O_2]+k_{l2}}\frac{\tilde{k}_{PO_1}[DHAA]}{1+(\tilde{k}_{PO_1}+\tilde{k}_{PO_y})[DHAA]}$$
$$L_p(1-\exp[-\kappa c_{DCA} l_{opt}]) - k_{PO_1 \cdot PO_x}[PO_1]$$
$$- k_{PO_1 \cdot BP_3}[TFA]^{m_{TFA}+1}[PO_1] - k_{PO_1 \cdot IM_2}[TFA]^{m_{TFA}}[PO_1]\Bigg),$$

$$\frac{d[PO_y]}{dz} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_l}\Bigg(\frac{[O_2]}{k_{l1}[O_2]+k_{l2}}\frac{\tilde{k}_{PO_y}[DHAA]}{1+(\tilde{k}_{PO_1}+\tilde{k}_{PO_y})[DHAA]}$$
$$L_p(1-\exp[-\kappa c_{DCA} l_{opt}]) - k_{PO_1 \cdot PO_x}[PO_y]\Bigg),$$

Table C.4
Estimated and literature parameters from our previous study on the photo reactor [23] that are used in the developed process model to identify the reaction network of the acid-catalyzed sequence.

Symbol	Unit	Value	Description
$\widetilde{k_l a}$	$\sqrt{1/\mathrm{cm\,min}}$	1.094	Effective mass transfer factor
$\tilde{k}_{\mathrm{PO_1}}$	$\mathrm{L\,mol^{-1}}$	7.130	Normalized kinetic rate constant
$\tilde{k}_{\mathrm{PO_y}}$	$\mathrm{L\,mol^{-1}}$	0.644	Normalized kinetic rate constant
l_{opt}	cm	0.178	Optical path length
C_0	–	1.02	Distribution parameter in drift flux model
$\tilde{L}_{\mathrm{p}}$	mol/(Ls%LED)	9.02×10^{-5}	Proportionality factor for incident photon flux
k_{l1}	–	0.641	Kinetic constant in quantum yield
k_{l2}	$\mathrm{mol\,L^{-1}}$	0.0119	Kinetic constant in quantum yield
κ	$\mathrm{L\,mol^{-1}\,cm^{-1}}$	12 841.98	Napierian absorption coefficient of DCA
τ_Δ	µs	33.2	Lifetime of singlet oxygen at −20 °C

$$\frac{\mathrm{d}[\mathrm{PO_x}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} k_{\mathrm{PO_1\cdot PO_x}}([\mathrm{PO_1}]+[\mathrm{PO_y}]),$$

$$\frac{\mathrm{d}[\mathrm{IM_{2c}}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} \big(k_{\mathrm{PO_1\cdot IM_2}}[\mathrm{TFA}]^{m_{\mathrm{TFA}}}[\mathrm{PO_1}] - k_{\mathrm{IM}_{2c}\cdot\mathrm{ART}}[\mathrm{TFA}][\mathrm{IM}_{2c}][\mathrm{O}_{2,1}] - k_{\mathrm{IM}_{2c}\cdot\mathrm{BP_5}}[\mathrm{TFA}][\mathrm{IM}_{2c}]\big),$$

$$\frac{\mathrm{d}[\mathrm{BP_3}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} k_{\mathrm{PO_1\cdot BP_3}}[\mathrm{TFA}]^{m_{\mathrm{TFA}}+1}[\mathrm{PO_1}],$$

$$\frac{\mathrm{d}[\mathrm{BP_5}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} k_{\mathrm{IM}_{2c}\cdot\mathrm{BP_5}}[\mathrm{TFA}][\mathrm{IM}_{2c}],$$

$$\frac{\mathrm{d}[\mathrm{ART}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} k_{\mathrm{IM}_{2c}\cdot\mathrm{ART}}[\mathrm{TFA}][\mathrm{IM}_{2c}][\mathrm{O}_{2,1}],$$

$$\frac{\mathrm{d}[\mathrm{O_2}]}{\mathrm{d}z} = \frac{\left(1-\frac{\beta}{C_0}\right)A}{\dot{V}_1} \Bigg(-\frac{[\mathrm{O_2}]}{k_{l1}[\mathrm{O_2}]+k_{l2}} \frac{(\tilde{k}_{\mathrm{PO_1}}+\tilde{k}_{\mathrm{PO_y}})[\mathrm{DHAA}]}{1+(\tilde{k}_{\mathrm{PO_1}}+\tilde{k}_{\mathrm{PO_y}})[\mathrm{DHAA}]} L_{\mathrm{p}}(1-\exp[-\kappa c_{\mathrm{DCA}} l_{\mathrm{opt}}]) + \widetilde{k_l a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}([\mathrm{O_2}]^\infty - [\mathrm{O_2}]) - k_{\mathrm{IM}_{2c}\cdot\mathrm{ART}}[\mathrm{TFA}][\mathrm{IM}_{2c}][\mathrm{O}_{2,1}]\Bigg),$$

$$\frac{\mathrm{d}x_{\mathrm{O_2}}}{\mathrm{d}z} = \frac{T}{p\dot{V}_{\mathrm{g}}}\left(-\bar{R}\frac{\beta}{C_0}A\widetilde{k_l a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}([\mathrm{O_2}]^\infty - [\mathrm{O_2}]) - \frac{p}{T}x_{\mathrm{O_2}}\frac{\mathrm{d}\dot{V}_{\mathrm{g}}}{\mathrm{d}z}\right),$$

$$\frac{\mathrm{d}\dot{V}_{\mathrm{g}}}{\mathrm{d}z} = -\frac{T}{p}\bar{R}\frac{\beta}{C_0}A\widetilde{k_l a}\sqrt{u_{\mathrm{g}}^{\mathrm{s}}}([\mathrm{O_2}]^\infty - [\mathrm{O_2}]).$$

A corresponding scheme is shown in Fig. 4. The following kinetic constants in the differential equations have been normalized to the lifetime of singlet oxygen:

$$\tilde{k}_i = k_i \tau_\Delta, i \in \{\mathrm{PO_1}, \mathrm{PO_y}\}. \tag{C.2}$$

The quantum yield of singlet oxygen is

$$\Phi_{^1\mathrm{O_2}} = \frac{[\mathrm{O_2}]}{k_{l1}[\mathrm{O_2}]+k_{l2}} \tag{C.3}$$

with k_{l1} and k_{l2} kinetic parameters that combine diverse rate constants of singlet oxygen generation. The local volumetric rate of photon absorption follows from Beer–Lambert law:

$$L_{\mathrm{p}}^{\mathrm{a}} = L_{\mathrm{p}}(1-\exp[-\kappa c_{\mathrm{DCA}} l_{\mathrm{opt}}]) \tag{C.4}$$

with $L_{\mathrm{p}} = \tilde{L}_{\mathrm{p}} P_{\mathrm{LED}}$ the local volumetric incident photon flux being proportional to the LED power percentage, κ the napierian absorption coefficient of the photosensitizer DCA, and l_{opt} the optical path length. Note that a differential equation for the molar oxygen fraction $x_{\mathrm{O_2}}$ was added. In contrast to the performed continuous experiments in this study with oxygen in the gas phase only ($x_{\mathrm{O_2}} = 1$), continuous experiments in our previous study were also run with a gas mixture of oxygen and nitrogen [23]. Please also note that the liquid volumetric flow $\dot{V}_1$ is predetermined by the fixed feed ratios given in Section 3.1, resulting in a jump condition at $z{=}z_{\mathrm{acid}}$. The corresponding initial conditions at the inlet of the photo reactor are

$$([\mathrm{DHAA}], [\mathrm{PO_1}], [\mathrm{PO_y}], [\mathrm{PO_x}], c_i, [\mathrm{O_2}], x_{\mathrm{O_2}}, \dot{V}_{\mathrm{g}})(z=0) = ([\mathrm{DHAA}]_0, 0, 0, 0, 0, [\mathrm{O_2}]^\infty, x_{\mathrm{O_2},0}, \dot{V}_{\mathrm{g},0}), \tag{C.5}$$

where the subscript 0 identifies initial conditions that can be set by the operator for each experimental run. The model parameters that have been determined in our previous publication are given in Table C.4.

Table D.5
Different calculation methods for the dissociation equilibrium of trifluoroacetic acid (TFA) on mass action law basis. Model A assumes full dissociation. Models B and C account for an incomplete dissociation due to the nonpolarity of the solvent toluene and protonable species, introducing a dissociation constant K. The main assumption is that the dissociation equilibrates fast compared with the decay of $\mathrm{PO_1}$, i.e., the equilibrium is in place before the actual reactions occur. Model B considers a simple equilibrium, where in model C it is assumed that all species evolving from and including DHAA, summarized in species S, intervene with the equilibration.

Diss.	Calculation of $\mathrm{H^+}$
A	$\mathrm{HTFA} \longrightarrow \mathrm{TFA^-} + \mathrm{H^+}$ $[\mathrm{H^+}] = [\mathrm{TFA}]_0$
B	$\mathrm{HTFA} \longleftrightarrow \mathrm{TFA^-} + \mathrm{H^+}$ $[\mathrm{H^+}] = -\frac{K}{2} + \sqrt{\frac{K^2}{4} + K[\mathrm{TFA}]_0}$
C	$\mathrm{HTFA} + \mathrm{S} \longleftrightarrow \mathrm{TFA^-} + \mathrm{SH^+}$ $[\mathrm{S}] = [\mathrm{DHAA}]_0, [\mathrm{SH^+}] = [\mathrm{H^+}]$ $[\mathrm{H^+}] = -\frac{K[\mathrm{DHAA}]_0}{2} + \sqrt{\frac{K^2[\mathrm{DHAA}]_0^2}{4} + K[\mathrm{TFA}]_0[\mathrm{DHAA}]_0}$

Appendix D. Summary of postulated kinetic models

The base model candidate is shown in Fig. D.14, and solely involves irreversible steps. $\mathrm{PO_1}$ conversion has two competing paths: It either reacts to the byproduct $\mathrm{BP_3}$ or to the intermediate $\mathrm{IM_2}$, with both reactions being acid-catalyzed. The series of formed species is summarized in the lumped species $\mathrm{IM_2}$, that represents the central enol intermediate; compare with Fig. 2. $\mathrm{IM_2}$ is the starting point for the reactions towards both ART, $\mathrm{BP_4}$, and $\mathrm{BP_5}$, where the latter species is introduced to compensate the low recoveries. $\mathrm{BP_4}$ is formed in an acid-catalyzed reaction step with the ejection of a water molecule. The ART formation is mapped in two steps.

Table D.6

Core set of proposed model candidates for the acid-catalyzed reaction network, corresponding to the superstructure in Fig. 8. Note that the font colors of the table headers also match the colors of the network characteristics in the superstructure illustration. The dissociation types, column 'Diss.', are given in Table D.5. '$\#x$' is the number of species (states) and 'q' is the number of model parameters to be estimated. The column 'Reaction orders' refers to the reaction orders regarding TFA from PO_1 to IM_2 and BP_3, respectively. All other reaction orders are unity. A check in the column 'IM_3' implies that IM_2 is participating and the orange direct route in Fig. 8 is inactive. If there is a check in any of the columns 'IM_2 prot.', 'BP_3 prot.' or 'BP_4 eq.', the corresponding colored route in Fig. 8 replaces the black route. In the columns 'Hydr. pos.' and '$BP_{5/6}$' the starting point for the corresponding reaction is given.

ID	$\#x$	q	Diss.	Reaction orders	IM_2 prot.	BP_3 prot.	BP_4 eq.	Hydr. pos.	$BP_{5/6}$ pos.	IM_3
1	9	7	A	1, 1	–	–	–	–	IM_2	✓
2	9	8	B	1, 1	–	–	–	–	IM_2	✓
3	9	8	C	1, 1	–	–	–	–	IM_2	✓
4	10	8	A	1, 1	✓	–	–	–	IM_2	–
5	11	9	A	1, 1	✓	–	–	–	IM_2	✓
6	10	9	A	1, 1	–	–	✓	–	IM_2	✓
7	12	11	A	1, 1	✓	✓	–	–	IM_2	✓
8	10	9	A	1, 1	–	–	–	IM_2	IM_2	✓
9	10	11	A	1, 1	–	–	✓	IM_2	IM_2	✓
10	13	13	A	1, 1	✓	✓	–	IM_2	IM_2	✓
11	13	13	A	1, 1	✓	✓	✓	–	IM_2	✓
12	13	15	A	1, 1	✓	✓	✓	IM_2	IM_2	✓
13	9	7	A	1, 2	–	–	–	–	IM_2	✓
14	9	8	A	$m_{IM_2}, 1$	–	–	–	–	IM_2	✓
15	9	8	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	–	IM_2	✓
16	9	8	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	–	IM_3	✓
17	9	9	B	$m_{IM_2}, m_{IM_2}+1$	–	–	–	–	IM_3	✓
18	9	9	C	$m_{IM_2}, m_{IM_2}+1$	–	–	–	–	IM_3	✓
19	11	10	A	$m_{IM_2}, m_{IM_2}+1$	✓	–	–	–	IM_2	✓
20	10	10	A	$m_{IM_2}, m_{IM_2}+1$	–	–	✓	–	IM_2	✓
21	10	10	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	IM_2	IM_2	✓
22	10	10	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	IM_3	IM_2	✓
23	10	10	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	IM_3	IM_3	✓
24	10	9	A	$m_{IM_2}, m_{IM_2}+1$	–	–	–	–	$IM_{2/3}$	✓
25	9	9	A	m_{IM_2}, m_{BP_3}	–	–	–	–	IM_2	✓
26	9	9	A	m_{IM_2}, m_{BP_3}	–	–	–	–	IM_3	✓
27	10	11	A	m_{IM_2}, m_{BP_3}	–	–	–	IM_2	IM_2	✓
28	10	11	A	m_{IM_2}, m_{BP_3}	–	–	✓	–	IM_2	✓
29	11	13	A	m_{IM_2}, m_{BP_3}	–	–	✓	IM_2	IM_2	✓
30	11	13	A	m_{IM_2}, m_{BP_3}	–	–	✓	IM_2	IM_3	✓

First, IM_2 reacts with oxygen to the intermediate IM_3, that physically represents an intermediate after the oxidation, Fig. 2. IM_3 is then further converted to ART in an acid-catalyzed reaction under the ejection of a water molecule. In the base network, the formed water does not affect the reaction kinetics, and the proton concentration, that is set to the initial TFA concentration, stays constant over the course of the reaction.

The first variants target the dissociation equilibrium of TFA in toluene. Despite being a very strong acid in water, dissociation of TFA in the nonpolar toluene is hampered. Additionally, protonated species such as PO_1 might interfere with the dissociation equilibrium. To account for the reduced dissociation, three different options, with the equations listed in Table D.5, are proposed. All of them hypothesize that the dissociation equilibrates significantly faster than the reactions with PO_1. Variant A assumes complete dissociation, variant B considers a dissociation equilibrium introducing the dissociation constant K, and variant C assumes that all species evolving from and including DHAA intervene in the dissociation equilibrium. The adjustment of the dissociation equilibrium is a consequence of the importance of a precise description of the acid-catalyzed conversion of PO_1 to yield accurate predictions for the products ART in the continuous and BP_4 in the batch experiments. To account for this importance, the reaction orders regarding proton concentration from PO_1 to IM_2 and from PO_1 to BP_3 are added to the set of model parameters considered during parameter estimation. Thus, the rate of reactions from the reaction scheme in Fig. D.14 are expressed as:

$$r_{PO_1 \cdot IM_2} = k_{PO_1 \cdot IM_2} [H^+]^{m_{IM_2}} [PO_1] \tag{D.1a}$$

$$r_{PO_1 \cdot BP_3} = k_{PO_1 \cdot BP_3} [H^+]^{m_{BP_3}} [PO_1] \tag{D.1b}$$

$$r_{PO_1 \cdot PO_x} = k_{PO_1 \cdot PO_x} [PO_1] \tag{D.1c}$$

$$r_{IM_2 \cdot IM_3} = k_{IM_2 \cdot IM_3} [IM_2][O_{2,l}] \tag{D.1d}$$

$$r_{IM_2 \cdot BP_4} = k_{IM_2 \cdot BP_4} [H^+][IM_2] \tag{D.1e}$$

$$r_{IM_2 \cdot BP_5} = k_{IM_2 \cdot BP_5} [H^+][IM_2] \tag{D.1f}$$

$$r_{IM_3 \cdot ART} = k_{IM_3 \cdot ART} [H^+][IM_3] \tag{D.1g}$$

, with reaction orders m_{BP_3} and m_{IM_2}, based on the rate law in Eq. (5). $[H^+]$ is given by the equations in Table D.5. In certain model candidates, the reaction orders are assumed to differ by unity, i.e., $m_{BP_3} = m_{IM_2} + 1$.

Apart from a modified dissociation mechanism and variable reaction orders, the base network was extended and altered to yield different model candidates, that comprehend reasonable reaction networks, with the goal of improving the model data fit. A superstructure of the developed model candidates is shown in Fig. 8, and a corresponding table containing the different model candidates in Table D.6.

In the base network, the initial protonation is considered to be rate-determining for the decay of PO_1. Alternatively, the rearrangement following the protonation might be rate-determining [58]. Here, the educt is first reversibly protonated and further converted to the product

Table D.7
Summary of the goodness of the fit, i.e. *RMSE* and AIC_c criterion, of the best parameter estimation runs for proposed core set of model candidates, ID according to Table D.6. The number of non-zero model parameters q is given in brackets in the respective column.

ID	q	*RMSE*	$RMSE_{conti}$	$RMSE_{batch}$	AIC_c
1	7 (6)	0.029	0.027	0.032	−4230
2	8 (7)	0.029	0.027	0.032	−4228
3	8 (7)	0.029	0.027	0.032	−4228
4	8 (6)	0.024	0.020	0.031	−4451
5	9 (7)	0.024	0.019	0.032	−4457
6	9 (8)	0.029	0.027	0.033	−4248
7	11 (11)	0.026	0.018	0.029	−4385
8	9 (7)	0.027	0.024	0.033	−4308
9	11 (8)	0.027	0.024	0.032	−4322
10	13 (12)	0.018	0.018	0.019	−4789
11	13 (13)	0.022	0.022	0.022	−4571
12	15 (14)	0.018	0.018	0.020	−4774
13	7 (6)	0.027	0.025	0.031	−4307
14	8 (7)	0.023	0.024	0.021	−4515
15	8 (7)	0.019	0.017	0.021	−4773
16	8 (8)	0.017	0.014	0.021	−4895
17	9 (9)	0.017	0.014	0.021	−4896
18	9 (9)	0.017	0.014	0.021	−4893
19	10 (10)	0.019	0.017	0.022	−4738
20	10 (9)	0.019	0.017	0.021	−4769
21	10 (9)	0.019	0.017	0.021	−4769
22	10 (9)	0.017	0.014	0.021	−4895
23	10 (10)	0.017	0.014	0.021	−4899
24	9 (8)	0.017	0.014	0.021	−4893
25	9 (8)	0.019	0.017	0.021	−4771
26	9 (9)	0.017	0.014	0.021	−4895
27	11 (10)	0.019	0.017	0.021	−4767
28	11 (10)	0.019	0.017	0.021	−4767
29	13 (12)	0.021	0.017	0.026	−4644
30	13 (13)	0.017	0.014	0.021	−4887

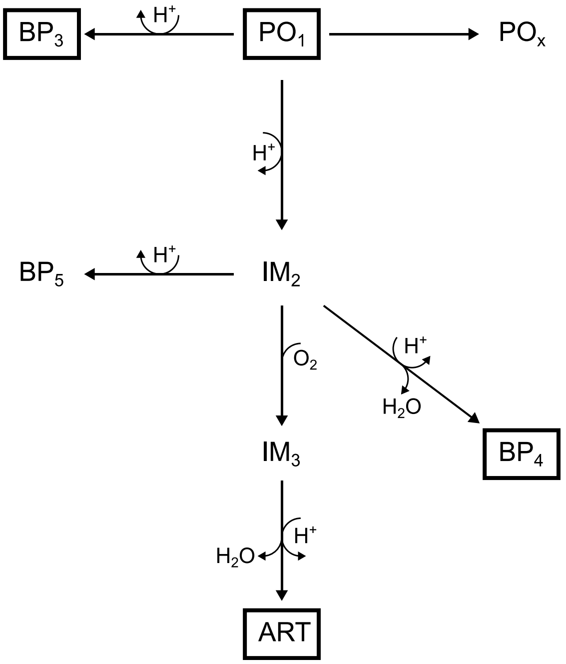


Fig. D.14. Scheme of base model candidate, i.e., model candidate 1 in Table D.6, for the reaction network of the acid-catalyzed sequence.

in a non-reversible acid-catalyzed step; see for example the path from PO_1 to IM_2 passing PO_1H^+ in Fig. 8.

The role of water, that can influence the acidity of the solution or change the reactivity of involved species by hydration, is taken into account in the species IM−hyd. Here, free water is trapped, where the equilibrium between IM_2 or IM_3 and IM−hyd shifts towards IM−hyd with increasing reaction progress, since more water becomes available. The equilibrium shifts cause the slow down of both the BP_4 and the ART formation.

In a further variant, the IM_3 species may not be present, that is, ART is directly formed from IM_2. The byproduct BP_5 can be derived either from IM_2 or from IM_3. Moreover, two byproducts, BP_5 and BP_6, can be formed from the two intermediates.

The results from the identification of the postulated model candidates is given in Table D.7.

Appendix E. Hatta number analysis

The Hatta number in the synthesis reactor is defined as [13]

$$Ha = \frac{\sqrt{D k_{IM_{2c}\cdot ART}[TFA][IM_{2c}]}}{\frac{\widetilde{k_1 a}\sqrt{u_g^s}}{a}} \tag{E.1}$$

with D=4.38 × $10^{-5}\,cm^2\,s^{-1}$ the diffusion coefficient of oxygen in toluene [59] and a the interfacial transfer area per unit volume. For obtaining an estimation about a maximum Hatta number, maximum values for TFA of 0.4 M and for IM_{2c} of 0.28 M are used. Note that the letter represents the maximum DHAA concentration, making it a significant overestimation of the maximum IM_{2c} concentration. The Hatta number becomes

$$Ha = \frac{\sqrt{60\cdot 4.38\times 10^{-5}\,cm^2\,min^{-1} 30.07\,L^2\,mol^{-2}\,min^{-1} 0.4\,mol\,L^{-1} 0.28\,mol\,L^{-1}}}{\frac{1.094\,\sqrt{/(cm\,min)}\sqrt{u_g^s}}{a}}$$

$$\approx \frac{0.094\,cm\,min^{-1}}{\frac{1.094\,\sqrt{/(cm\,min)}\sqrt{u_g^s}}{a}}.$$

If the longest synthesis reactor (6 m length, 0.16 cm diameter) was filled with gas only, the specific area would be a=25 cm^{-1}. The maximum Hatta number then becomes

$$Ha \approx \frac{0.094\,cm\,min^{-1}}{\frac{1.094\,\sqrt{/(cm\,min)}\sqrt{u_g^s}}{25\,cm^{-1}}} \approx \frac{1}{\sqrt{u_g^s}} 2.15\,\sqrt{cm/min}$$

The process is transfer-controlled for Ha>3 and kinetically controlled for Ha<0.3. Hence, with the condition Ha>1 for partial transfer limitations, the superficial gas velocity would need to be approximately lower than 4 $cm\,min^{-1}$ for partial transfer limitations. However, gas velocities are substantially higher, i.e., initial superficial gas velocities approximately vary from 80 $cm\,min^{-1}$ to 400 $cm\,min^{-1}$ at 0.08 cm inner tube diameter, Table A.3. This translates to a range of 20 $cm\,min^{-1}$ to 100 $cm\,min^{-1}$ at 0.16 cm inner tube diameter. The process is therefore running in the regime controlled by the chemical kinetics.

Data availability

The data presented in the manuscript are openly available in the KITopen repository at https://doi.org/10.35097/2vwx8arcenkf1jpc.

References

[1] World Health Organization, WHO Briefing on Malaria Treatment Guidelines and Artemisinin Monotherapies, Technical Report, 2006, p. 28.

[2] N. Ma, Z. Zhang, F. Liao, T. Jiang, Y. Tu, The birth of artemisinin, Pharmacol. Ther. 216 (2020) 107658, http://dx.doi.org/10.1016/j.pharmthera.2020.107658.

[3] J.L. Siqueira-Neto, K.J. Wicht, K. Chibale, J.N. Burrows, D.A. Fidock, E.A. Winzeler, Antimalarial drug discovery: Progress and approaches, Nat. Rev. Drug Discov. 22 (10) (2023) 807–826, http://dx.doi.org/10.1038/s41573-023-00772-9.

[4] Z.-w. Zeng, D. Chen, L. Chen, B. He, Y. Li, A comprehensive overview of artemisinin and its derivatives as anticancer agents, Eur. J. Med. Chem. 247 (2023) 115000, http://dx.doi.org/10.1016/j.ejmech.2022.115000.

[5] D.H. Cheong, D.W. Tan, F.W. Wong, T. Tran, Anti-malarial drug, artemisinin and its derivatives for the treatment of respiratory diseases, Pharmacol. Res. 158 (2020) 104901, http://dx.doi.org/10.1016/j.phrs.2020.104901.

[6] M. Peplow, Synthetic biologys first malaria drug meets market resistance, Nature 530 (2016) 389–390, http://dx.doi.org/10.1038/530390a.

[7] C.J. Paddon, P.J. Westfall, D.J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M.D. Leavell, A. Tai, A. Main, D. Eng, D.R. Polichuk, K.H. Teoh, D.W. Reed, T. Treynor, J. Lenihan, H. Jiang, M. Fleck, S. Bajad, G. Dang, D. Dengrove, D. Diola, G. Dorin, K.W. Ellens, S. Fickes, J. Galazzo, S.P. Gaucher, T. Geistlinger, R. Henry, M. Hepp, T. Horning, T. Iqbal, L. Kizer, B. Lieu, D. Melis, N. Moss, R. Regentin, S. Secrest, H. Tsuruta, R. Vazquez, L.F. Westblade, L. Xu, M. Yu, Y. Zhang, L. Zhao, J. Lievense, P.S. Covello, J.D. Keasling, K.K. Reiling, N.S. Renninger, J.D. Newman, High-level semi-synthetic production of the potent antimalarial artemisinin, Nature 496 (7446) (2013) 528–532, http://dx.doi.org/10.1038/nature12051.

[8] D.-K. Ro, E.M. Paradise, M. Ouellet, K.J. Fisher, K.L. Newman, J.M. Ndungu, K.A. Ho, R.A. Eachus, T.S. Ham, J. Kirby, M.C.Y. Chang, S.T. Withers, Y. Shiba, R. Sarpong, J.D. Keasling, Production of the antimalarial drug precursor artemisinic acid in engineered yeast, Nature 440 (7086) (2006) 940–943, http://dx.doi.org/10.1038/nature04640.

[9] R.J. Roth, N. Acton, A simple conversion of artemisinic acid into artemisinin, J. Nat. Prod. 52 (5) (1989) 1183–1185, http://dx.doi.org/10.1021/np50065a050.

[10] N. Acton, R.J. Roth, On the conversion of dihydroartemisinic acid into artemisinin, J. Org. Chem. 57 (13) (1992) 3610–3614, http://dx.doi.org/10.1021/jo00039a020.

[11] R.J. Roth, N. Acton, A facile semisynthesis of the antimalarial drug Qinghaosu, J. Chem. Educ. 68 (7) (1991) 612, http://dx.doi.org/10.1021/ed068p612.

[12] M.B. Plutschack, B. Pieber, K. Gilmore, P.H. Seeberger, The Hitchhiker's guide to flow chemistry, Chem. Rev. 117 (18) (2017) 11796–11893, http://dx.doi.org/10.1021/acs.chemrev.7b00183.

[13] Y. Su, N.J.W. Straathof, V. Hessel, T. Noël, Photochemical transformations accelerated in continuous-flow reactors: Basic concepts and applications, Chem. Eur. J. 20 (34) (2014) 10562–10589, http://dx.doi.org/10.1002/chem.201400283.

[14] M. Lancel, B. Tambosco, Z. Amara, Photochemical routes to artemisinin, in: Photochemistry: Volume 51, Royal Society of Chemistry, 2023, http://dx.doi.org/10.1039/BK9781837672301-00301.

[15] F. Lévesque, P.H. Seeberger, Continuous-flow synthesis of the anti-malaria drug artemisinin, Angew. Chem. Int. Ed. 51 (7) (2012) 1706–1709, http://dx.doi.org/10.1002/anie.201107446.

[16] D. Kopetzki, F. Lévesque, P.H. Seeberger, A continuous-flow process for the synthesis of artemisinin, Chem. Eur. J. 19 (17) (2013) 5450–5456, http://dx.doi.org/10.1002/chem.201204558.

[17] S. Triemer, K. Gilmore, G.T. Vu, P.H. Seeberger, A. Seidel-Morgenstern, Literally green chemical synthesis of artemisinin from plant extracts, Angew. Chem. Int. Ed. 57 (19) (2018) 5525–5528, http://dx.doi.org/10.1002/anie.201801424.

[18] K. Varela, F.K. Yoshimoto, Syntheses of deuterium-labeled dihydroartemisinic acid (DHAA) isotopologues and mechanistic studies focused on elucidating the conversion of DHAA to artemisinin, Org. Biomol. Chem. 22 (43) (2024) 8527–8550, http://dx.doi.org/10.1039/D4OB00777H.

[19] K. Varela, H.D. Arman, F.K. Yoshimoto, Synthesis of [15,15,15- $^2H_3$]-dihydroartemisinic acid and isotope studies support a mixed mechanism in the endoperoxide formation to artemisinin, J. Nat. Prod. 84 (7) (2021) 1967–1984, http://dx.doi.org/10.1021/acs.jnatprod.1c00246.

[20] K. Varela, H.D. Arman, F.K. Yoshimoto, Synthesis of [3,3- $^2H_2$]-dihydroartemisinic acid to measure the rate of nonenzymatic conversion of dihydroartemisinic acid to artemisinin, J. Nat. Prod. 83 (1) (2020) 66–78, http://dx.doi.org/10.1021/acs.jnatprod.9b00686.

[21] E.N. Pistikopoulos, A. Barbosa-Povoa, J.H. Lee, R. Misener, A. Mitsos, G.V. Reklaitis, V. Venkatasubramanian, F. You, R. Gani, Process systems engineering – The generation next? Comput. Chem. Eng. 147 (2021) http://dx.doi.org/10.1016/j.compchemeng.2021.107252.

[22] P. Sinner, S. Daume, C. Herwig, J. Kager, Usage of digital twins along a typical process development cycle, in: C. Herwig, R. Pörtner, J. Möller (Eds.), Digital Twins, vol. 176, Springer International Publishing, Cham, 2020, pp. 71–96, http://dx.doi.org/10.1007/10_2020_149.

[23] S. Triemer, M. Schulze, B. Wriedt, R. Schenkendorf, D. Ziegenbalg, U. Krewer, A. Seidel-Morgenstern, Kinetic analysis of the partial synthesis of artemisinin: Photooxygenation to the intermediate hydroperoxide, J. Flow Chem. 11 (3) (2021) 641–659, http://dx.doi.org/10.1007/s41981-021-00181-2.

[24] D. Langary, Z. Nikoloski, Inference of chemical reaction networks based on concentration profiles using an optimization framework, Chaos: An Interdiscip. J. Nonlinear Sci. 29 (11) (2019) 113121, http://dx.doi.org/10.1063/1.5120598.

[25] D.G. Blackmond, Reaction progress kinetic analysis: A powerful methodology for mechanistic studies of complex catalytic reactions, Angew. Chem. Int. Ed. 44 (28) (2005) 4302–4320, http://dx.doi.org/10.1002/anie.200462544.

[26] L.P. De Oliveira, D. Hudebine, D. Guillaume, J.J. Verstraete, A review of kinetic modeling methodologies for complex processes, Oil Gas Sci. Technol. – Rev. d'IFP Energies Nouv. 71 (3) (2016) 45, http://dx.doi.org/10.2516/ogst/2016011.

[27] C.J. Taylor, A. Pomberger, K.C. Felton, R. Grainger, M. Barecka, T.W. Chamberlain, R.A. Bourne, C.N. Johnson, A.A. Lapkin, A brief introduction to chemical reaction optimization, Chem. Rev. 123 (6) (2023) 3089–3126, http://dx.doi.org/10.1021/acs.chemrev.2c00798.

[28] S. Triemer, Reactive Transformation of Extraction Byproducts: Enhanced Production of the Antimalarial Artemisinin (Ph.D. thesis), Otto-von-Guericke-Universität Magdeburg, Magdeburg, 2023, http://dx.doi.org/10.25673/103427.

[29] K.P. Burnham, D.R. Anderson, K.P. Burnham, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, second ed., Springer, New York, 2002, http://dx.doi.org/10.1007/b97636.

[30] R.K. Haynes, S.C. Vonwiller, Catalysed oxygenation of allylic hydroperoxides derived from qinghao (artemisinic) acid. Conversion of qinghao acid into dehydroginghaosu (artemisitene) and qinghaosu (artemisinin), J. Chem. Soc. Chem. Commun. (6) (1990) 451, http://dx.doi.org/10.1039/c39900000451.

[31] S.C. Vonwiller, J.A. Warner, S.T. Mann, R.K. Haynes, Copper(II) trifluoromethanesulfonate-induced cleavage oxygenation of allylic hydroperoxides derived from qinghao acid in the synthesis of Qinghaosu derivatives: Evidence for the intermediacy of enols, J. Am. Chem. Soc. 117 (45) (1995) 11098–11105, http://dx.doi.org/10.1021/ja00150a009.

[32] R.K. Haynes, S.C. Vonwiller, From qinghao, marvelous herb of antiquity, to the antimalarial trioxane qinghaosuand some remarkable new chemistry, Acc. Chem. Res. 30 (2) (1997) 73–79, http://dx.doi.org/10.1021/ar950058w.

[33] L.-K. Sy, G.D. Brown, R. Haynes, A novel endoperoxide and related sesquiterpenes from Artemisia annua which are possibly derived from allylic hydroperoxides, Tetrahedron 54 (17) (1998) 4345–4356, http://dx.doi.org/10.1016/S0040-4020(98)00148-3.

[34] L.-K. Sy, K.-S. Ngo, G.D. Brown, Biomimetic synthesis of arteannuin h and the 3,2-rearrangement of allylic hydroperoxides, Tetrahedron 55 (52) (1999) 15127–15140, http://dx.doi.org/10.1016/S0040-4020(99)00987-4.

[35] L.-K. Sy, G.D. Brown, The mechanism of the spontaneous autoxidation of dihydroartemisinic acid, Tetrahedron 58 (5) (2002) 897–908, http://dx.doi.org/10.1016/S0040-4020(01)01193-0.

[36] G.D. Brown, L.-K. Sy, In vivo transformations of dihydroartemisinic acid in Artemisia annua plants, Tetrahedron 60 (5) (2004) 1139–1159, http://dx.doi.org/10.1016/j.tet.2003.11.070.

[37] L. Feng, Y. Wang, S. Yuan, K.-Y. Wang, J.-L. Li, G.S. Day, D. Qiu, L. Cheng, W.-M. Chen, S.T. Madrahimov, H.-C. Zhou, Porphyrinic metal–organic frameworks installed with brønsted acid sites for efficient tandem semisynthesis of artemisinin, ACS Catal. 9 (6) (2019) 5111–5118, http://dx.doi.org/10.1021/acscatal.8b04960.

[38] M.A. Corsello, N.K. Garg, Synthetic chemistry fuels interdisciplinary approaches to the production of artemisinin, Nat. Prod. Rep. 32 (3) (2015) 359–366, http://dx.doi.org/10.1039/C4NP00113C.

[39] S.H. Kung, S. Lund, A. Murarka, D. McPhee, C.J. Paddon, Approaches and recent developments for the commercial production of semi-synthetic artemisinin, Front. Plant Sci. 9 (2018) 87, http://dx.doi.org/10.3389/fpls.2018.00087.

[40] A. Seidel-Morgenstern, Analysis and experimental investigation of catalytic membrane reactors, in: K. Sundmacher, A. Kienle, A. Seidel-Morgenstern (Eds.), Integrated Chemical Processes, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, 2005, pp. 359–389, http://dx.doi.org/10.1002/3527605738.ch12.

[41] O. Levenspiel, Chemical Reaction Engineering, third ed., Wiley, New York, 1999.

[42] A. Roibu, S. Kuhn, Modeling of photochemical processes in continuous-flow reactors, in: T. Noël (Ed.), Photochemical Processes in Continuous-Flow Reactors, World Scientific, New Jersey and London and Singapore, 2017, pp. 69–96, http://dx.doi.org/10.1142/9781786342195_0003.

[43] C. Yao, Y. Zhao, H. Ma, Y. Liu, Q. Zhao, G. Chen, Two-phase flow and mass transfer in microchannels: A review from local mechanism to global models, Chem. Eng. Sci. 229 (2021) 116017, http://dx.doi.org/10.1016/j.ces.2020.116017.

[44] M. Ishii, T. Hibiki, Thermo-Fluid Dynamics of Two-Phase Flow, Springer New York, NY, 2006, http://dx.doi.org/10.1007/978-0-387-29187-1.

[45] G.M. Nirmal, T.F. Leary, A. Ramachandran, Mass transfer dynamics in the dissolution of Taylor bubbles, Soft Matter 15 (13) (2019) 2746–2756, http://dx.doi.org/10.1039/C8SM01144C.

[46] C. Vandu, H. Liu, R. Krishna, Mass transfer from Taylor bubbles rising in single capillaries, Chem. Eng. Sci. 60 (22) (2005) 6430–6437, http://dx.doi.org/10.1016/j.ces.2005.01.037.

[47] X. Wu, Z. Deng, J. Yan, Z. Zhang, F. Zhang, Z. Zhang, Experimental investigation on the solubility of oxygen in toluene and acetic acid, Ind. Eng. Chem. Res. 53 (23) (2014) 9932–9937, http://dx.doi.org/10.1021/ie5014772.

[48] M. Schulze, Model-Based Process Identification for Pharmaceutical Synthesis (Ph.D. thesis), Karlsruher Institut für Technologie (KIT), 2024, http://dx.doi.org/10.5445/IR/1000167542.

[49] R.J. Littel, G.F. Versteeg, W.P.M. Van Swaaij, Diffusivity measurements in some organic solvents by a gas-liquid diaphragm cell, J. Chem. Eng. Data 37 (1) (1992) 42–45, http://dx.doi.org/10.1021/je00005a014.

[50] S. Haase, D.Y. Murzin, T. Salmi, Review on hydrodynamics and mass transfer in minichannel wall reactors with gas–liquid Taylor flow, Chem. Eng. Res. Des. 113 (2016) 304–329, http://dx.doi.org/10.1016/j.cherd.2016.06.017.

[51] E. Walter, L. Pronzato, Identification of parametric models from experimental data, Communications and Control Engineering, Springer, London, 1997.

[52] J. Bezanson, A. Edelman, S. Karpinski, V.B. Shah, Julia: A fresh approach to numerical computing, SIAM Rev. 59 (1) (2017) 65–98, http://dx.doi.org/10.1137/141000671.

[53] E. Roduner, Understanding catalysis, Chem. Soc. Rev. 43 (24) (2014) 8226–8239, http://dx.doi.org/10.1039/C4CS00210E.

[54] A. Schöniger, T. Wöhling, L. Samaniego, W. Nowak, Model selection on solid ground: Rigorous comparison of nine ways to evaluate Bayesian model evidence, Water Resour. Res. 50 (12) (2014) 9484–9513, http://dx.doi.org/10.1002/2014WR016062.

[55] C. Alamillo-Ferrer, G. Hutchinson, J. Burés, Mechanistic interpretation of orders in catalyst greater than one, Nat. Rev. Chem. 7 (1) (2022) 26–34, http://dx.doi.org/10.1038/s41570-022-00447-w.

[56] J. Burés, What is the order of a reaction? Top. Catal. 60 (8) (2017) 631–633, http://dx.doi.org/10.1007/s11244-017-0735-y.

[57] R. Gupta, D. Fletcher, B. Haynes, Taylor flow in microchannels: A review of experimental and computational work, J. Comput. Multiph. Flows 2 (1) (2010) 1–31, http://dx.doi.org/10.1260/1757-482X.2.1.1.

[58] M. Beller, A. Renken, R.A. van Santen (Eds.), Catalysis: From Principles to Applications, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2012.

[59] A. Schumpe, P. Luehring, Oxygen diffusivities in organic liquids at 293.2 K, J. Chem. Eng. Data 35 (1) (1990) 24–25, http://dx.doi.org/10.1021/je00059a007.