A Self-Driving Closed-Loop Workflow for Data-Efficient Kinetic Modeling and Optimization of the Aldol Reaction

Xiao Li[1], Pierre-Louis Lagueux-Tremblay[2], Alan Cherney[2], Abhishek Paul[1], William Walker[1]
[1]*U.S. Process Engineering, CMC Synthetics Platform, Sanofi, 350 Water St, Cambridge, Massachusetts 02141, United States* [2] *U.S. Process Chemistry, CMC Synthetics Platform, Sanofi, 350 Water St, Cambridge, Massachusetts 02141, United States*

Correspondence to: Yasser Jangjou Yasser.Jangjou@sanofi.com

1. Introduction

A self-driving, closed-loop workflow is developed to autonomously iterate between experimentation and modeling, maximizing information gain from minimal data. The workflow integrates automated time-course data collection using a Chemspeed platform equipped with online HPLC, mechanistically informed kinetic model construction, and model-driven experimental design. At each iteration, the model evaluates model uncertainty and discrimination power to propose the next most informative experiment, enabling rapid refinement of kinetic parameters and efficient exploration of reaction space. This information-driven, data-efficient framework provides a systematic pathway for accelerating reaction optimization while deepening mechanistic understanding.

As a demonstration of this approach, the workflow is applied to the direct (classical) aldol reaction, one of the most fundamental carbon–carbon bond-forming transformations in organic synthesis. From a sustainability perspective, the direct aldol reaction is attractive due to its operational simplicity, low cost, and minimal environmental impact, often requiring only mixing and stirring under mild conditions.[1] However, these advantages are offset by persistent challenges, including poor chemo- and stereoselectivity, low isolated yields, and limited product stability. Despite extensive historical study of aldol chemistry and its mechanism, as reviewed by Mestres and others, the literature reveals a notable lack of systematic optimization and quantitative kinetic data for direct aldol reactions.[1,2] In contrast, alternative approaches such as the Mukaiyama aldol reaction offer improved selectivity but typically rely on less environmentally benign reagents and conditions.[3]

2. Method

All experiments were executed using a Chemspeed robotic platform interfaced with a large language model (LLM) through Chemspyd,[4] which serves as the command translation layer. For the initial experiment, the user provides a textual description of the intended experimental procedure. The LLM interprets this description and generates Chemspyd commands, which are then executed directly on the Chemspeed platform.

The automated workflow for reaction setup includes the following steps:

1. Weighing and stock solution preparation: Solid reagents are weighed automatically, and stock solutions for the reaction are prepared on-deck.
2. Reaction initiation: Starting materials, along with any necessary reaction aids, are added to initiate the reaction under controlled conditions.
3. Time-course data collection: Reaction progress is monitored in real time using online HPLC. The resulting data are automatically processed by the algorithm.

The self-driving component of the workflow continuously integrates experimental data with kinetic modeling (Figure 1). Mechanistically informed models of the reaction are constructed, with uncertainties in kinetic parameters explicitly quantified. Based on these models, subsequent experiments are autonomously proposed with the dual objectives of reducing model uncertainty and maximizing reaction yield. Each new experiment is translated by the LLM into executable Chemspeed commands, completing the closed-loop cycle. This approach enables data-efficient exploration of reaction space, leveraging real-time experimental feedback to iteratively refine kinetic understanding and optimize reaction conditions.

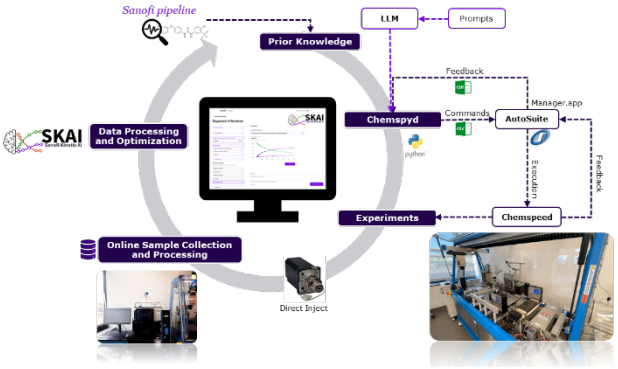
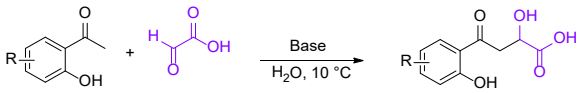


Fig. 1: Overview of closed loop workflow

3. Results

The reaction we selected is the aldol addition between an acetophenol and glyoxylic acid to form α-hydroxy acid; a useful class of molecule, that can be utilized for downstream synthesis of pharmaceuticals (Scheme 1).[5] However, it was observed that subsequent undesired cascade reactions, such as dehydration, cyclization, or subsequent aldol additions, were difficult to prevent due to the lability of the hydroxyl group in the product.



Scheme 1: Aldol reaction between acetophenols and glyoxylic acid to yield α-hydroxy acids (AHAs) under basic aqueous conditions.

Preliminary studies of the selected aldol reaction indicate that product stability is the primary challenge. A series of reactions were conducted using various bases and temperatures on an ReactAll platform,[6] capable of running five reactions in parallel at a 5 mL scale (Figure 2a and 2b). Among the conditions tested, the use of triethylamine (TEA) at 10 °C yielded a relatively stable aldol product, with impurity levels remaining around 0.5 % after three days (Figure 2b), demonstrating the feasibility of achieving stable product under mild conditions.

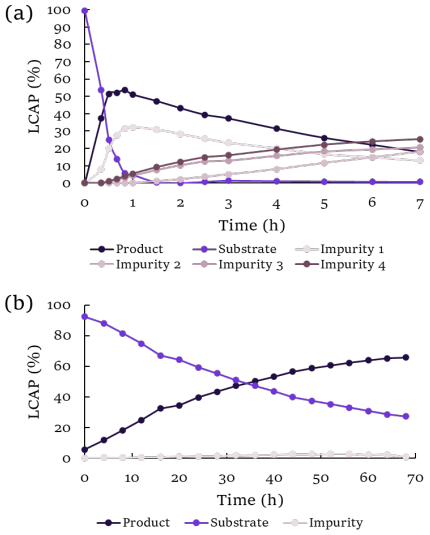


Figure 2. (a) Liquid Chromatography Area Percent (LCAP) of the aldol reaction with KOH at 20 °C; (b) LCAP of the aldol reaction with TEA at 10 °C.

These findings validate the reaction setup and provide a benchmark for further optimization. The self-driving closed-loop workflow enables subsequent experimentation on a significantly smaller 2 mL scale, combining automated reaction execution, real-time online HPLC monitoring, and model-driven experiment design. By leveraging this approach, the workflow is poised to systematically refine reaction conditions to further improve yield and product stability while minimizing experimental resource use. Additional time-course data collected through the self-driving closed-loop platform are currently in progress and will inform ongoing model refinement and optimization.

Acknowledgments

The authors are grateful to Shawn Walker and Jason Tedrow from the CMC Synthetics Platform for their support.

References

[1] Mestres, R. A Green Look at the Aldol Reaction. Green Chem. 2004, 6, 583–603.
[2] Perrin, C. L.; Chang, K.-L. The Complete Mechanism of an Aldol Condensation. J. Org. Chem. 2016, 81, 5631–5635.
[3] Matsuo, J.-i.; Murakami, M. The Mukaiyama Aldol Reaction: 40 Years of Continuous Development. Angew. Chem. Int. Ed. 2013, 52, 9109–9118.
[4] Seifrid, M.; *et al. Chemspyd: An Open-Source Python Interface for Chemspeed Robotic Chemistry and Materials Platforms.* Digital Discovery 2024, *3*, 1319–1326.

[5] Bianchi, M., et al. Gastric anti-secretory, anti-ulcer and cytoprotective properties of substituted (E)-4-phenyl- and heteroaryl-4-oxo-2-butenoic acids. Eur. J. Med. Chem. 1988, 23, 45-52.
[6] ReactAll. ReactALL - Technobis Crystallization Systems