Flow-based Methods in Chemical Peptide and Protein Synthesis

Kevin Schiefelbein and Nina Hartrampf*

Abstract: Flow chemistry has emerged as a powerful method for on-demand chemical synthesis and modification of peptides and proteins. Herein, we discuss the characteristics of flow chemistry and how they are applied to various aspects of peptide chemistry. We highlight recent advances in automated flow-based peptide synthesis, which extend the length of peptides routinely accessible to single-domain proteins and allow for the collection of time-resolved synthesis data. Applications of this data for the prediction of synthesis outcome and the potential for the development of more sustainable synthesis methods are also discussed. Finally, we will review solution-phase approaches, including flow-based ligation strategies and peptide cyclization. Throughout this review, the current challenges and potential future developments are highlighted.

Keywords: Flow chemistry · Ligation · Peptides and proteins · Solid-phase peptide synthesis

1. Introduction

Peptides, proteins, and derivatives thereof can be used as therapeutics and provide tools to investigate biological mechanisms at the molecular level, and their efficient production is of key interest for academia and the pharmaceutical industry. Peptides and proteins can be successfully obtained through biological expression or chemical synthesis. Chemical synthesis provides access to a plethora of substitutions and modifications, including non-canonical amino acids, post-translational modifications (PTMs), artificial backbone motifs, and macrocyclization for peptide and protein design. In most cases, these reactions are performed in batch, however, flow-based methods are attracting renewed interest given the advantages, such as precise, reproducible control of reaction conditions, and minimized side-reactions. Flow chemistry is a continuous process whereby reagents are transferred and reacted within tubing or a designated reactor. As a result, stoichiometry is controlled by flow rate and reagent concentrations, and reaction time is defined by residence time in the reactor.[1] In general, flow-reactions can be performed in solution by continuously mixing reagents in flow, or *via* a solid-phase approach using immobilized reagents or catalysts in a fixed-bed reactor.[2] Continuous flow offers four key advantages for peptide chemistry: (1) Reactive or transient intermediates can be generated and immediately consumed under highly controlled reaction conditions (flash chemistry),[3] (2) continuous flow ensures minimal exposure of reactor contents to reactive by-products, (3) energy transfer is improved, and (4) time-resolved reaction monitoring is achieved by in-line analysis.[4]

2. Solid-phase Peptide and Protein Chemistry in Flow

Solid-phase peptide synthesis (SPPS) serves as the foundation of modern chemical peptide and protein production.[5] Briefly, an N^α-protected amino acid is linked to an insoluble polymer support (resin) and the peptide sequence is generated through iterative N^α-deprotection and amino acid coupling steps. A global deprotection and cleavage of the peptidyl-resin releases the soluble peptide. Despite decades of optimization, peptides longer than 50 amino acids are difficult to synthesize due to sequence-dependent aggregation on resin and statistical accumulation of by-products.[6] In

contrast to traditional synthesis methods, flow-based SPPS offers high control over reaction conditions, and in-line analysis gives valuable information on individual coupling and deprotection steps. We have recently reviewed the general principle of flow-based SPPS and set the method into a historical context,[4] thus this review will focus on recent developments.

2.1 *Automated Fast-flow Peptide Synthesis (AFPS)*

The Pentelute group developed a manual and an automated 'fast-flow' SPPS protocol, centered around a heated activation loop and reactor (Fig. 1).[7,8] By automating this process, significant improvements with respect to fluid handling and reproducibility were achieved. The coupling reagents and protected amino acids are combined in the heated loop under highly controlled micro-flow conditions, generating an active ester, which is then delivered to a tube reactor containing the SPPS-resin. In contrast to conventional peptide synthesizers, AFPS also provides real-time information on the fluorenylmethyloxycarbonyl (Fmoc) deprotection steps by UV-Vis detection of dibenzofulvene. downstream of the resin bed. Then, the integral, height, and width of the generated UV-Vis signals can indicate the efficiency of each deprotection to pinpoint the steps in which peptide aggregation occurs. Furthermore, the swelling and shrinking of the peptidyl-resin during synthesis can be measured to monitor coupling steps (swelling), deprotection steps (shrinking), and aggregation (shrinking) using a Variable-Bed Reactor technology developed by Seeberger and coworkers.[9]

Flow-based SPPS was recently optimized to enable the rapid synthesis of large and complex peptides and proteins.[10] A systematic screen of conditions and optimization of machine parameters led to significantly improved coupling efficiency. A three-step optimization approach was followed to refine the synthesis recipe: first, we developed a set of conditions generally applicable independent of the chemical properties of the coupled buildng block. Then, deprotection data obtained from in-line UV-Vis analysis for all 20 canonical amino acids was statistically analyzed to develop an amino acid-dependent synthesis recipe. Finally, the coupling of the epimerization-prone residues, cysteine and histidine, was analyzed to determine the temperature- and time-dependence of

Correspondence: Prof. N. Hartrampf, E-mail: nina.hartrampf@chem.uzh.ch
Department of Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
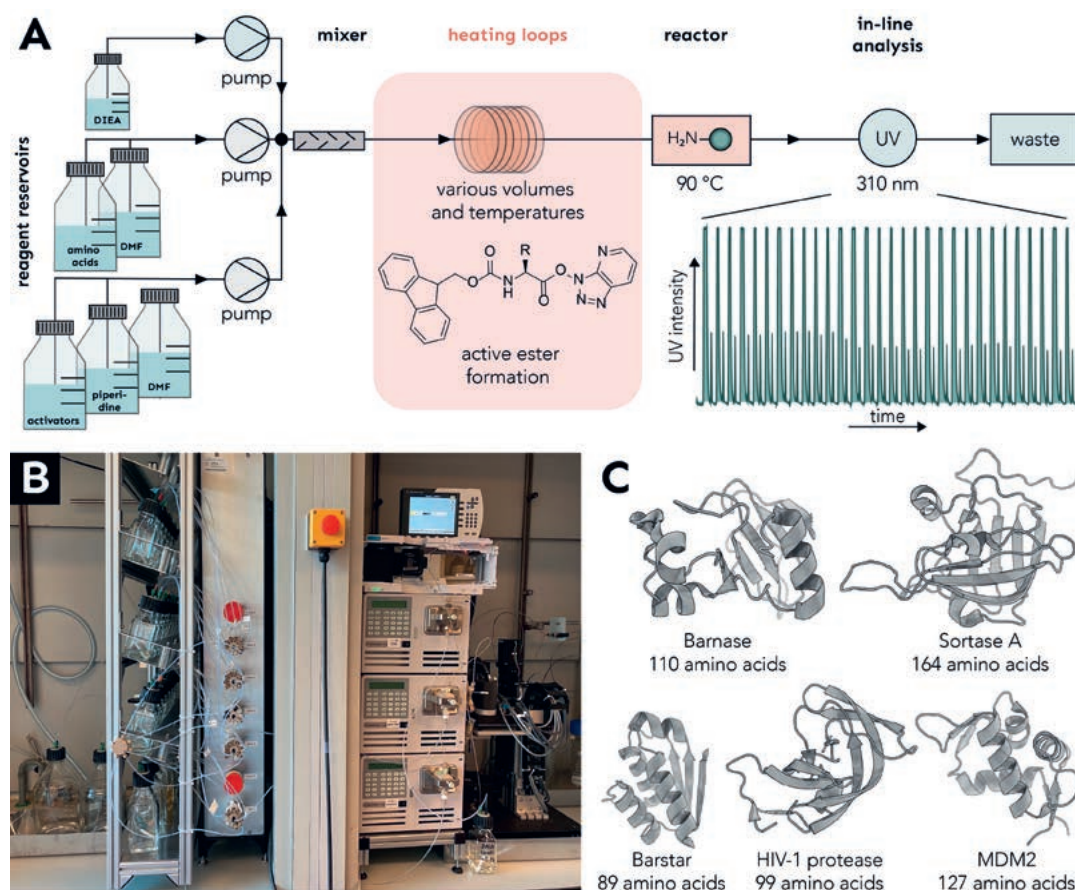
Fig. 1. Automated Fast Flow Synthesis (AFPS). (A) Schematic representation of the AFPS and its operating principle. The reagents are delivered by HPLC pumps to the mixer followed by amino acid activation in a heated loop before coupling occurs in the reactor at high temperatures. The in-line UV-analysis is performed at 310 nm to obtain time-dependent traces for the synthesis as exemplarily shown. (B) Photograph of the AFPS system in the Hartrampf lab at the University of Zurich. (C) Synthesized proteins using the AFPS system. These have been purified and folded showing activities comparable to their expressed counterparts (pdb: 2ZA4, 1BTA, 3HBO, 3G03, 2KID).

epimerization. Using this optimized general synthesis recipe, long peptides of up to 164 amino acids were successfully produced at a rate of 25 couplings per hour. To validate this approach, the structure and activity of the folded synthetic proteins were compared to the respective recombinantly-expressed proteins. Recently, this optimized AFPS technology was applied in the synthesis of immobilized proteins for library screening,[11] new GLP-1 derivatives,[12] and computationally-predicted cell-penetrating peptides.[13]

2.2 Deep Learning for Peptide Synthesis Prediction

Automated in-line data collection produces highly reproducible data sets, which can then be used for computational analysis and deep learning approaches. The sequence-dependence of peptide and protein synthesis was investigated by Pentelute, Gómez-Bombarelli and coworkers using the time-dependent UV-traces from AFPS platforms.[14] For this approach, the in-line UV-traces of 35427 Fmoc-removal steps were analyzed to evaluate amino acid-specific coupling efficiencies. A computational model was developed to map the structural representations of amino acids and peptide sequences to the experimental synthesis parameters, to ultimately predict the outcome of specific deprotection reactions. Through this deep learning approach, sequence-specific aggregation events and Fmoc-deprotection efficiencies can be accurately predicted for synthesis optimization (Fig. 2). In the future, deep learning may enable real-time optimization of peptide synthesis in flow and if combined with on-chip flow-SPPS, such as described by Wang and coworkers,[15] rapid small-scale optimization can be achieved.

2.3 Green Peptide Synthesis

There is a pressing need to develop greener and more sustainable alternatives to every aspect of SPPS, especially with increasing interest in peptide-based therapeutics. Standard protocols employ excess protected amino acids and toxic coupling reagents as well

as hazardous solvents (*e.g. N,N*-dimethylformamide, DMF), resulting in the accumulation of toxic waste and poor atom economy.[16]

One advantage of flow chemistry is that reagents pass through the packed bed in a narrow, high-concentration band, which improves coupling efficiency and theoretically requires less reagent than batch-SPPS.[17] Fülöp and coworkers demonstrated the flow synthesis of unmodified and *N*-methylated peptides with only 1.5 equivalents of amino acids and coupling reagents in a flow-based system.[18,19] For batch-SPPS, much research has been devoted to the search for new SPPS-compatible solvents to replace polar, aprotic DMF. However, despite the recommendation of numerous green solvent alternatives, none of them have been tested in flow-based SPPS. A more substantial change to current synthesis methods could arise from adapting SPPS to the 'ideal' green solvent water. Advances in water-compatible protecting groups and activators have recently been reported,[20,21] and warrant investigation in flow-SPPS.

3. Solution-phase Approaches for the Synthesis of Peptides and Proteins in Flow

As an alternative to SPPS, amide bond formation for the synthesis of peptides and proteins can also be achieved in solution. Here, we present methodologies for stepwise synthesis of small peptides in continuous flow, as well as flow-based cyclization methods and peptide ligation for construction of larger compounds.

3.1 Solution-phase Methods for Peptide Bond Formation in Flow

Recent reviews on amide bond formation reveal an underrepresentation of flow chemistry methods.[22] This is surprising given amide bond formation in peptide synthesis requires mild and fine-tuned conditions to activate amino acids without epimerization or undesired loss of protecting groups – features that flow chemistry is perfectly suited for. The field
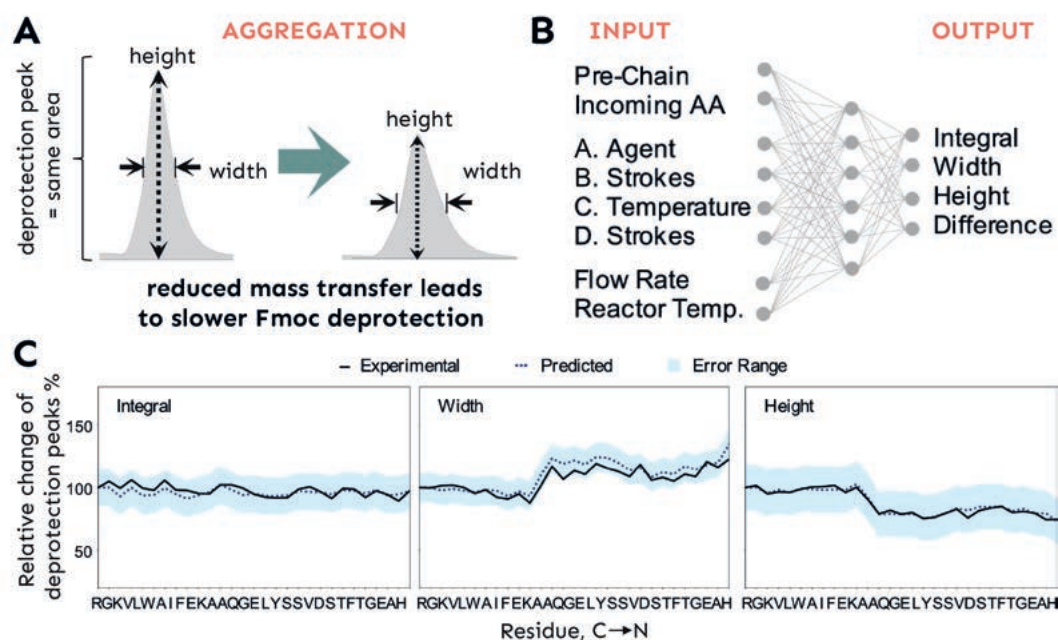
Fig. 2. Deep learning using UV-Vis Fmoc-removal data from an AFPS allows for the accurate prediction of aggregation. (A) On-resin aggregation of the peptides causes broadening of the Fmoc-deprotection peak in the UV-analysis. (B) Schematic representation of the machine learning parameters. The input describes the screened parameters for the coupling reaction of the incoming amino acids, and the output refers to the UV-analysis of the subsequent Fmoc-deprotection step. (C) The peak width, height, and integral for each Fmoc deprotection is analyzed to indicate the efficiency of each coupling and deprotection cycle. The occurrence of aggregation is characterized by an increased peak width and decreased peak height. It is shown that the predicted outcome concurs with the experimentally observed data.

of 'flash chemistry' leverages these highly controlled reaction conditions to generate reactive or transient intermediates that are immediately quenched by their reaction partner before decomposing or undergoing side reactions.[3] Takahashi and coworkers showed a strong and rapid activation of amino acids using triphosgene in microflow (0.5 s for activation) to form dipeptides in less than ten seconds at ambient temperature.[23] Notably, only little epimerization was observed due to the fast trapping of the highly reactive symmetric anhydride by the amine. In addition, Fuse *et al.* reported on the use of mixed carbonic anhydrides for amide bond formation in flow.[24]

Another approach to generate a highly reactive amino acid intermediate *in situ* was shown by Kappe and coworkers.[25] Beginning with acyl hydrazides, acyl azides were generated inline and then reacted with *C*-terminally protected amino acids to afford dipeptides. Alternatively, Jamison and coworkers used a photochemical nitrone rearrangement for amide bond formation using amino aldehyde and nitrone starting materials.[26]

As an alternative to the *in situ* activation of the amino acids in flow, *N*-carboxy anhydrides (NCAs) can be used as stable, yet reactive, amino acid derivatives for amide bond formation. However, NCAs must be utilized under highly controlled reaction conditions to avoid polymerization, a requirement fulfilled with flow chemistry (Fig. 3). Ley and coworkers used Cbz-protected NCAs to synthesize a hexamer using automated flow methodology.[27] This was then expanded by Blacker and coworkers who applied unprotected NCAs in the flow synthesis of peptide dimers and trimers with a low by-product profile.[28]

A catalytic approach to solution-phase amide bond formation is the use of hydrogen chloride-assisted activation.[29] In this method, an acyl *N*-methylimidazolium cation is utilized to promote amide bond formation in flow. This approach has enabled the synthesis of *N*-methylated dipeptides in solution with reduced reaction times compared to standard conditions.[29]

Immobilized reagents and catalysts can also be used in flow for carboxylic acid activation. To date, two catch-and-release

strategies using polymer-bound hydroxybenzotriazole[30] and carbodiimides[31] have been reported. In these methodologies, the carboxylic acid is 'caught' and activated by the resin-bound catalyst and then undergoes nucleophilic attack by the incoming amine, affording the desired amide bond and releasing the product from the resin.
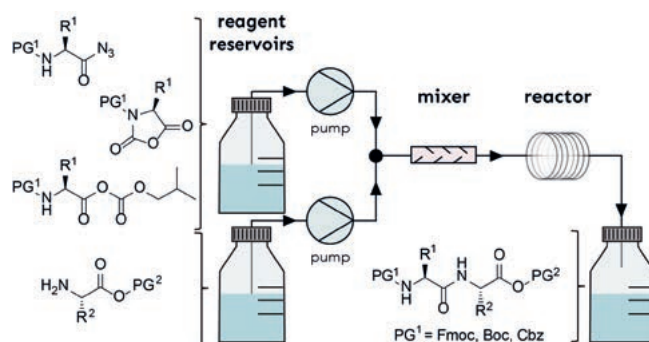

Fig. 3. Highly activated amino acids (acyl azides, NCAs, mixed anhydrides) with applications in flow-based solution-phase synthesis of small peptides.

3.2 *Flow-based Peptide Ligation and Bioconjugation*

Beyond the coupling of individual amino acids, a range of methodologies – including flow-based approaches – were developed for the syntheses of proteins from unprotected peptides. Payne and coworkers successfully performed Native Chemical Ligation (NCL) in flow for the synthesis of the HIV-inhibitor Enfuvirtide.[32] Using this approach, the ligation time (32 min) was significantly reduced in comparison to batch processes (overnight). Payne and coworkers also applied in flow photochemical cysteine desulfurization to obtain the desired HIV-inhibitor. Furthermore, Pentelute and coworkers demonstrated an enzymatic ligation employing Sortase A to synthesize protein bioconjugates in flow.[33]

Another strategy to covalently link two peptides makes use of bioorthogonal chemical handles, such as azides and alkynes. Hatit *et. al.* used copper(I) salts, formed by deliberate erosion of copper pipes, to catalyze a 1,3-dipolar cycloaddition between alkyne- and azide-modified peptides in flow.[34]

3.3 *Flow-based Peptide Cyclization Methods*

Cyclic peptides are becoming increasingly important in drug development due to their conformational rigidity as well as their enhanced resistance to proteolysis. However, in the synthesis of cyclic peptides, the desired cyclization reaction is in competition with dimerization and oligomerization. As a result, these reactions are typically performed at low concentrations to favor the intramolecular cyclization reaction. Flow chemistry for peptide cyclization has proven advantageous through the precise control of reaction parameters to prevent dimer and oligomer formation. Fuse and coworkers have reported a microflow-system for the cyclization of synthetic peptides by an intramolecular photochemical reaction of a pentapeptide bearing a *C*-terminal, photo activatable 5-bromo-7-nitroindoline (Bni)-protecting group and free *N*-terminal amine, affording the head-to-tail cyclized peptide.[35] Furthermore, microfluidic technology has been applied in peptide cyclization using NCL methods.[36]

4. Summary

Flow chemistry offers unique opportunities to complement existing techniques in peptide and protein chemistry, which we have only recently started to explore. In the near future, advancements in flow chemistry and its automation are highly anticipated. In particular, in-line analysis and autonomous systems coupled with deep learning are expected to further improve the synthesis and modification of peptides and proteins in flow. In this review, we highlighted the recent developments in solid-phase and solution-phase peptide synthesis in the context of flow chemistry, and recent advances in flow-based ligation and peptide modification. Although several challenges, such as sustainability of peptide synthesis and site-selectivity of peptide and protein modifications remain unsolved, flow chemistry is poised to significantly expand the areas of peptide chemistry, medicinal chemistry and chemical biology.

[1] K. F. Jensen, R. Becker, M. M. E. Delville, M. Fekete, F. Fülöp, T. Glasnov, T. A. Hamlin, R. K. Harmel, C. O. Kappe, K. Koch, N. E. Leadbeater, H. Löwe, A. Macchi, I. M. Mándity, P. Nieuwland, S. B. Ötvös, P. Plouffe, D. Roberge, F. Rutjes, J. Yoshida, 'Fundamentals', De Gruyter, **2014**, https://doi.org/10.1515/9783110289169.

[2] S. B. Ötvös, I. M. Mándity, F. Fülöp, *ChemSusChem* **2012**, *5*, 266, https://doi.org/10.1002/cssc.201100332.

[3] J. Yoshida, Y. Takahashi, A. Nagaki, *Chem. Commun.* **2013**, *49*, 9896, https://doi.org/10.1039/C3CC44709J.

[4] Z. P. Gates, N. Hartrampf, *Pept. Sci.* **2020**, *112*, e24198, https://doi.org/10.1002/pep2.24198.

[5] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149, https://doi.org/10.1021/ja00897a025.

[6] S. B. H. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338, https://doi.org/10.1039/B700141J.

[7] M. D. Simon, Ph.D. Thesis Massachusetts Institute of Technology, Cambridge, MA, USA, **2017**.

[8] A. J. Mijalis, D. A. Thomas, M. D. Simon, A. Adamo, R. Beaumont, K. F. Jensen, B. L. Pentelute, *Nat. Chem. Biol.* **2017**, *13*, 464, https://doi.org/10.1038/nchembio.2318.

[9] E. T. Sletten, M. Nuño, D. Guthrie, P. H. Seeberger, *Chem. Commun.* **2019**, *55*, 14598, https://doi.org/10.1039/C9CC08421E.

[10] N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen, B. L. Pentelute, *Science* **2020**, *368*, 980, https://doi.org/10.1126/science.abb2491.

[11] A. J. Quartararo, Z. P. Gates, B. A. Somsen, N. Hartrampf, X. Ye, A. Shimada, Y. Kajihara, C. Ottmann, B. L. Pentelute, *Nat. Commun.* **2020**, *11*, 3183, https://doi.org/10.1038/s41467-020-16920-3.

[12] C. K. Longwell, S. Hanna, N. Hartrampf, R. A. P. Sperberg, P.-S. Huang, B. L. Pentelute, J. R. Cochran, *ACS Chem. Biol.* **2021**, *16*, 58, https://doi.org/10.1021/acschembio.0c00722.

[13] C. K. Schissel, S. Mohapatra, J. M. Wolfe, C. M. Fadzen, K. Bellovoda, C.-L. Wu, J. A. Wood, A. B. Malmberg, A. Loas, R. Gomez-Bombarelli, B. L. Pentelute, *bioRxiv* **2020**, 2020.04.10.036566, https://doi.org/10.1101/2020.04.10.036566.

[14] S. Mohapatra, N. Hartrampf, M. Poskus, A. Loas, R. Gómez-Bombarelli, B. L. Pentelute, *ACS Cent. Sci.* **2020**, *6*, 2277, https://doi.org/10.1021/acscentsci.0c00979.

[15] W. Wang, Y. Huang, J. Liu, Y. Xie, R. Zhao, S. Xiong, G. Liu, Y. Chen, H. Ma, *Lab. Chip* **2011**, *11*, 929, https://doi.org/10.1039/c0lc00542h.

[16] K. E. Schwieter, J. N. Johnston, *J. Am. Chem. Soc.* **2016**, *138*, 14160, https://doi.org/10.1021/jacs.6b08663.

[17] E. Bayer, *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 113, https://doi.org/10.1002/anie.199101133.

[18] I. M. Mándity, B. Olasz, S. B. Ötvös, F. Fülöp, *ChemSusChem* **2014**, *7*, 3172, https://doi.org/10.1002/cssc.201402436.

[19] A. Szloszár, I. M. Mándity, F. Fülöp, *J. Flow Chem.* **2018**, *8*, 21, https://doi.org/10.1007/s41981-018-0002-9.

[20] S. Knauer, C. Uth, N. Koch, R. Meusinger, O. Avrutina, H. Kolmar, *Angew. Chem. Int. Ed.* **2020**, *59*, 12984, https://doi.org/10.1002/anie.202003676.

[21] Q. Wang, Y. Wang, M. Kurosu, *Org. Lett.* **2012**, *14*, 3372, https://doi.org/10.1021/ol3013556.

[22] R. M. de Figueiredo, J.-S. Suppo, J.-M. Campagne, *Chem. Rev.* **2016**, *116*, 12029, https://doi.org/10.1021/acs.chemrev.6b00237.

[23] S. Fuse, Y. Mifune, T. Takahashi, *Angew. Chem. Int. Ed.* **2014**, *53*, 851, https://doi.org/10.1002/anie.201307987.

[24] S. Fuse, K. Masuda, Y. Otake, H. Nakamura, *Chem. – Eur. J.* **2019**, *25*, 15091, https://doi.org/10.1002/chem.201903531.

[25] A. Mata, U. Weigl, O. Flögel, P. Baur, C. A. Hone, C. O. Kappe, *React. Chem. Eng.* **2020**, *5*, 645, https://doi.org/10.1039/D0RE00034E.

[26] Y. Zhang, M. L. Blackman, A. B. Leduc, T. F. Jamison, *Angew. Chem. Int. Ed.* **2013**, *52*, 4251, https://doi.org/10.1002/anie.201300504.

[27] K. Rahbek Knudsen, M. Ladlow, Z. Bandpey, S. Ley, *J. Flow Chem.* **2015**, *4*, 18, https://doi.org/10.1556/JFC-D-13-00033.

[28] K. E. Jolley, W. Nye, C. González Niño, N. Kapur, A. Rabion, K. Rossen, A. J. Blacker, *Org. Process Res. Dev.* **2017**, *21*, 1557, https://doi.org/10.1021/acs.oprd.7b00214.

[29] Y. Otake, Y. Shibata, Y. Hayashi, S. Kawauchi, H. Nakamura, S. Fuse, *Angew. Chem. Int. Ed.* **2020**, *59*, 12925, https://doi.org/10.1002/anie.202002106.

[30] I. R. Baxendale, S. V. Ley, C. D. Smith, G. K. Tranmer, *Chem. Commun.* **2006**, 4835, https://doi.org/10.1039/b612197g.

[31] C. Dankers, J. Tadros, D. G. Harman, J. R. Aldrich-Wright, T. V. Nguyen, C. P. Gordon, *ACS Comb. Sci.* **2020**, *22*, 255, https://doi.org/10.1021/acscombsci.0c00001.

[32] T. S. Chisholm, D. Clayton, L. J. Dowman, J. Sayers, R. J. Payne, *J. Am. Chem. Soc.* **2018**, *140*, 9020, https://doi.org/10.1021/jacs.8b03115.

[33] R. L. Policarpo, H. Kang, X. Liao, A. E. Rabideau, M. D. Simon, B. L. Pentelute, *Angew. Chem.* **2014**, *126*, 9357, https://doi.org/10.1002/ange.201403582.

[34] M. Z. C. Hatit, L. F. Reichenbach, J. M. Tobin, F. Vilela, G. A. Burley, A. J. B. Watson, *Nat. Commun.* **2018**, *9*, 4021, https://doi.org/10.1038/s41467-018-06551-0.

[35] Y. Mifune, H. Nakamura, S. Fuse, *Org. Biomol. Chem.* **2016**, *14*, 11244, https://doi.org/10.1039/C6OB02391F.

[36] N. Ollivier, T. Toupy, R. C. Hartkoorn, R. Desmet, J.-C. M. Monbaliu, O. Melnyk, *Nat. Commun.* **2018**, *9*, 2847, https://doi.org/10.1038/s41467-018-05264-8.