Quantitative Monitoring of Solvent Concentration and Moisture Content in Flow Synthesis

Yusuke Hattori *1 *Yasuyuki Suzuki* *1

In this work, with the aim of quantitatively monitoring solvent concentration and moisture content in synthesis processes, a micro-flow cell and regression model of near-infrared (NIR) spectra were developed and utilized for monitoring peptide synthesis. The flow cell was designed to have sufficient resistance to pressure and sufficient displacement efficiency of the reaction fluid. The regression models employed the partial least squares regression method. The quantified concentration of each solvent and moisture content were equivalent to the theoretical values with low standard deviations. The standard deviation of moisture content was 10 ppm. Even a tiny amount of water can have a critical adverse impact on some chemical reactions. Therefore, strict control of moisture content is required. The proposed method for monitoring solvent concentration and moisture content can be simultaneously conducted with monitoring of other reaction factors, such as reaction products. The present results indicate the utility of this method for in-line spectroscopic monitoring of synthesis processes.

INTRODUCTION

The synthesis of many compounds, including pharmaceuticals, is currently carried out by batch synthesis, in which solvents and raw materials are added to a reaction vessel and the resulting product is purified after the reaction is complete. Batch synthesis makes it possible to synthesize compounds with complex structures through multistep reaction processes, but a significant amount of time and resources is required for the separation and purification of intermediates obtained at each stage, as well as for analysis using methods such as high-performance liquid chromatography.

In contrast, flow synthesis enables the continuous execution of multistep reactions by passing a flow of raw materials through narrow channels, and carrying out mixing and reactions within the channels[(1)]. Furthermore, by performing liquid chromatography or spectroscopic measurements directly within the channels, the analysis time can be significantly reduced compared to batch synthesis. The excellent real-time capability of spectroscopic measurements makes them particularly suitable for analyzing the continuously progressing reactions in flow synthesis.

Another issue with batch synthesis is that the use of large-capacity reaction vessels results in a delay before reaction conditions such as temperature reach their set values[(2)(3)]. In addition, the amount of energy required increases proportionally with the vessel volume[(2)(3)]. In contrast, in flow synthesis, where the reaction solution is transported through narrow channels, the surface area per unit volume of liquid is large, resulting in a high heat exchange rate. Therefore, reaction conditions such as temperature can be efficiently controlled with less energy and without response delay.

However, due to the high heat exchange rate, flow synthesis is also more susceptible to environmental changes, and requires strict environmental control as well as rigorous monitoring and regulation of the reaction solution. In particular, the precipitation of solute molecules caused by a drop in temperature can lead to blockage of the flow channels,

*1 Life Research & Development Department, Innovation Center, Marketing Headquarters

and therefore requires careful attention. Blockage of the channels necessitates the suspension of flow synthesis and considerable effort for cleaning and recovery. Therefore, in addition to cleaning the flow paths and managing temperature, it is essential to thoroughly monitor and eliminate factors that could lead to the formation or precipitation of impurities.

Another factor that can cause precipitation is a change in the solvent. For example, the introduction of a poor solvent for a solute or changes in the composition of a solvent mixture can reduce the solubility of solute molecules. In addition, it has been reported that in lithiation reactions using organolithium compounds, the intrusion of water leads to the formation and precipitation of lithium hydroxide, resulting in blockage of the flow channels[4][5]. Therefore, monitoring solvent concentration is also important; however, doing so is difficult, and no appropriate monitoring method exists.

The molecules of typical organic solvents are generally low in reactivity and chemically stable. They do not absorb ultraviolet or visible light, making detection and quantification using these wavelengths difficult. As a result, vibrational spectroscopy methods such as infrared spectroscopy, Raman spectroscopy, and near-infrared (NIR) spectroscopy are effective[6]. However, although infrared spectroscopy is capable of measuring very strong absorptions, a drawback is the absorption becomes too large and saturates for analytes present at high concentrations such as solvents. Raman spectroscopy, on the other hand, requires an excitation laser to be irradiated at the measurement point, which limits the measurement region. Therefore, in this study, a method for monitoring solvent concentration using NIR spectroscopy was investigated, and a compact transmission measurement cell (micro-flow cell) suitable for real-time measurement in flow synthesis was developed. To achieve high pressure resistance and displacement efficiency, a flow cell with an optical path length of 2 mm was designed. Partial least squares regression (PLSR), a multivariate analysis method, was used for data analysis to quantitatively predict and monitor solvent concentrations.

EXPERIMENT

Samples

Dipeptide synthesis using the liquid-phase tagging method (STag-PS method)[7] was used as a model reaction. As organic solvents, 4-methyltetrahydropyran (MTHP; density at 25°C: 0.86 g/cm^3; molecular weight: 100.16), *N,N*-dimethylformamide (DMF; density at 25°C: 0.95 g/cm^3; molecular weight: 73.09), and *N,N*-dimethylacetamide (DMA; density at 25°C: 0.94 g/cm^3; molecular weight: 87.12) were used. The water used was ultrapure.

In the STag-PS method, an amino acid modified at the C-terminus with an STag bearing a hydrophobic long-chain alkyl group and a trimethylsilyl group (PeptiStar, Osaka) was synthesized (H-Arg(Pbf)-O(-STag); molecular weight: 1201.94)[8]. The amino acid used for condensation with this was Fmoc-Oic-OH (molecular weight: 391.47; Ambeed, IL, USA). Ethyl 2-cyano-2-((dimethyliminio)(morpholino) methyloxyimino)acetate hexafluorophosphate (COMU; molecular weight: 428.27; Merck, NJ, USA) and *N,N*-diisopropylethylamine (DIEA; molecular weight: 129.25; Tokyo Chemical Industry) were used as condensing agents.

Method

NIR Spectral Measurement System For spectral measurements, a Fourier transform near-infrared spectrometer (FTNIR-L1-025-2TE, ARCoptix, Neuchâtel, Switzerland) was used with a halogen lamp (ARCLight-NIR, 20 W, ARCoptix) as the light source. A low-OH silica optical fiber (Z37L01, Thorlabs, NJ, USA), which exhibits low absorption loss due to OH groups and has a numerical aperture of 0.22 and a core diameter of 550 µm, was connected to the emission port of the light source (SMA-905). The other end of the fiber was connected to the flow cell used for spectroscopic measurement. The transmitted light from the flow cell was connected to the spectrometer's entrance port (NA 0.25, SMA-905) using an identical optical fiber. A flow cell equipped with a quartz cell having an optical path length of 2 mm was developed (Figure 1). The cell volume was 16 µL (channel cross-section: 2 mm × 2 mm; channel length: 4 mm). The flow cell and the flow channel were connected by flat-bottom fittings (1/4-28 UNF, IDEX, IL, USA). Pressure resistance was experimentally measured up to 6 MPa.

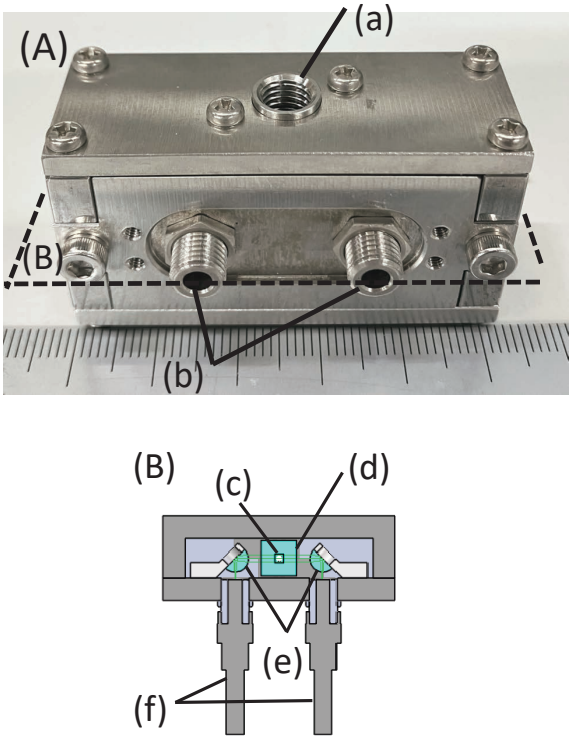

Figure 1 (A) Image of the flow cell: (a) 1/4-28 UNF fluidic port, (b) SMA905 optical fiber connectors; (B) Horizontal cross-sectional view of the central part of the flow cell with optical fibers connected (corresponding to the dashed line in (A)): (c) measurement channel (2 mm × 2 mm), (d) quartz cell, (e) sapphire hemispherical lens, (f) optical fiber

NIR Spectral Measurement and Regression Analysis A 1/8-inch PTFE tube (inner diameter: 1.59 mm, outer diameter: 3.2 mm) was connected to the aforementioned flow cell using flat-bottom fittings. Using a plunger pump (UI-22-110, FLOM, Tokyo), the test solution was delivered from the reaction vessel (50 mL flask) to the flow cell, during which spectral measurements were conducted. The test solution delivered to the flow cell was then returned to the vessel through a similarly connected PTFE tube, allowing it to circulate continuously.

Test solutions were prepared in 25 different compositions to cover the concentration ranges of MTHP, DMF, DMA, and water as shown in Table 1. For each composition, spectral measurements were performed three times. Spectra were acquired with a wavenumber resolution of 8 cm^{-1} and 32 scans per measurement. All measurements were conducted at a temperature of 40°C. Regression analysis of solvent concentration and moisture content was performed using PLSR.

Table 1 Concentration range of each solvent

	Conc. Range	
	mol/L	Vol%
MTHP	3.46～7.73	40～90
DMF	1.29～5.23	10～40
DMA	0～2.38	0～22
Water	0～1.09	0～2.0

The measured transmission spectra (S_{obs}) were converted into absorption spectra (A) using the spectrum of a reference sample (DMF:MTHP = 38.28:61.72 [v/v]) (S_{ref}) according to the following equation, after which second-derivative processing was applied.

$$A = -\log_{10}\frac{S_{obs}}{S_{ref}}$$

Second-derivative processing was performed using the Savitzky–Golay method, with a window size of 11 points and a polynomial order of 3.

Prediction of Solvent Concentration in Batch Synthesis of Peptides A total of 18 mL of solution was prepared by adding 1.5 mmol of the amino acid (H-Arg(Pbf)-O-STag) used as the starting material to a solvent mixture (DMF:MTHP = 10:90 (v/v)) in a reaction vessel (50-mL pear-shaped flask). This solution was circulated through the flow cell as described above. The pump flow rate was set to 9.0 mL/min. Using a syringe pump (ULTRA-P, Harvard, MA, USA), Fmoc-Oic-OH dissolved in DMA (0.4 M), DIEA mixed with MTHP (1.76 M), and COMU dissolved in DMF (0.4 M) were added to the reaction vessel at flow rates of 0.2 mL/min, 0.1 mL/min, and 0.2 mL/min, respectively, and the synthesis reaction was carried out. Under these conditions, spectral measurements were conducted using the flow cell, and solvent concentration during synthesis was predicted using the regression model obtained by PLSR.

Prediction of Solvent Concentration in Flow Synthesis of Peptides H-Arg(Pbf)-O(-STag), Fmoc-Oic-OH, DIEA, and COMU, prepared in the same manner as for batch synthesis, were delivered into a flow synthesis device[8] using a plunger pump, and flow synthesis was carried out. Following activation of Fmoc-Oic-OH by DIEA and COMU (5 s), a coupling reaction with H-Arg(Pbf)-O(-STag) was performed (60 s at 40°C). The flow cell was placed immediately downstream of this, where spectral measurements were taken. Solvent concentration after flow synthesis was predicted using the regression model obtained by PLSR.

RESULTS AND DISCUSSION

Construction of Regression Models by PLSR

Figure 2 shows the second-derivative NIR spectra used for constructing the regression models. The wavenumber ranges of the spectra used in PLSR were optimized individually for each solvent and for water. For water, characteristic peaks were observed at 5200 cm^{-1} and 5300 cm^{-1}, both of which are attributed to combination bands of OH stretching and bending vibrations[9]. This indicates the presence of both strongly hydrogen-bonded water molecules (peak at 5200 cm^{-1}) and weakly hydrogen-bonded water molecules (peak at 5300 cm^{-1}) in the reaction solution. Regression analysis including both peaks was performed to analyze water quantitatively. Absorption peaks due to water are known to shift with temperature[10]. Since this shift arises from the temperature dependence of hydrogen bonding strength, all measurements in this study were conducted under constant temperature conditions (40°C).

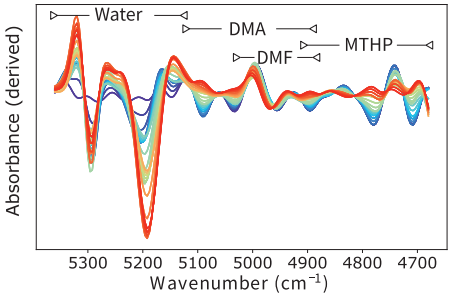

Figure 2 NIR spectra used for regression analysis of solvent concentration by PLSR. The ranges indicated by arrows represent the wavenumber regions used for regression analysis of each compound. Color differences correspond to differences in composition

The root-mean-squared error (RMSE), coefficient of determination (R^2), and number of latent variables (LV) are summarized in Table 2 as the results of regression analysis of solvent concentration using PLSR. The R^2 value was greater than 0.99 in all cases, with calibration errors of approximately 1% for MTHP, DMF, and water, and about 2% for DMA. The number of LVs was determined with consideration of both model error and prediction accuracy. Three LVs were used for all solvents except DMA, for which six were used. Although the number of LVs for DMA was relatively high, it was based

on the results of solvent concentration prediction for batch and flow synthesis.

Table 2 Results of regression analysis by PLSR

	RMSE, mol/L	R^2	LV
MTHP	0.0396	0.999	3
DMF	0.0743	0.996	3
DMA	0.0720	0.994	6
Water	0.0116	0.998	3

Figure 3 shows the second-derivative NIR spectra and the corresponding PLSR regression vectors for each solvent (MTHP, DMF, and DMA). For MTHP, characteristic peaks were observed near 4900 cm^{-1}, 4860 cm^{-1}, and in the range of 4800 to 4780 cm^{-1}. Although the 4900 cm^{-1} peak overlaps with a DMA peak, the other peaks, particularly the one in the 4800 to 4780 cm^{-1} range, contributed strongly to the regression. For DMF, there was a distinct peak near 5000 cm^{-1} and a broad peak between 4930 and 4920 cm^{-1}, both of which contributed significantly to the regression. For DMA, characteristic peaks were observed near 5020 cm^{-1}, 4970 cm^{-1}, and 4900 cm^{-1}. These results indicate that the concentration of each solvent can be predicted from its corresponding spectrum without being influenced by the presence of the other solvents.

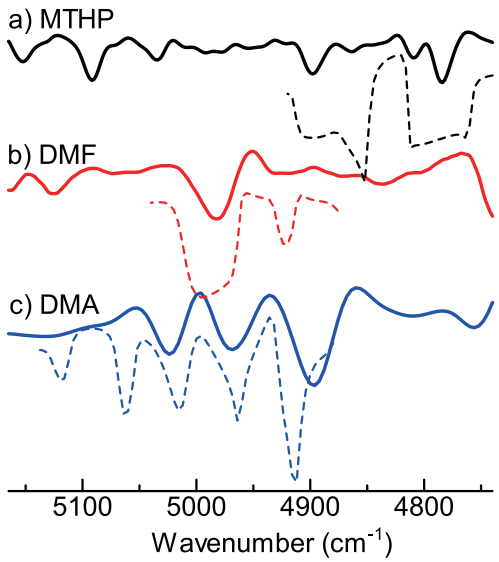

Figure 3 Second-derivative NIR spectra (solid lines) and regression vectors (dashed lines) for each solvent

Prediction of Solvent Concentration in Batch Synthesis

In the batch synthesis, the STag starting material dissolved in a mixed solvent of DMF:MTHP = 10:90 was first added to the reaction vessel. COMU, Fmoc-Oic-OH, and DIEA were each dissolved in DMF, DMA, and MTHP, respectively, and added to the reaction vessel at constant flow rates using a syringe pump. The coupling reaction between the amino acids progressed over time, and the solvent concentrations also changed, so the peaks originating from the solvents varied accordingly over time.

Two peaks attributed to water molecules were present in the range of 5160 to 5300 cm^{-1} before the reaction began. Since water was not added as a reagent, it is presumed that residual water remained from the synthesis and purification of the STag starting material. In addition, the reagents added via syringe pump did not contain water, so these peaks decreased over time. The other peaks were derived from MTHP. Since the amount of MTHP added was small, its concentration decreased from the initial level, and the peak intensity decreased accordingly.

Figure 4(A) shows the results of applying the regression model constructed by PLSR to the spectra obtained during batch synthesis to predict solvent concentrations. For comparison, theoretical concentrations of MTHP, DMF, and DMA were calculated. These theoretical concentrations were determined based on the initial concentrations, the delivery rates of the syringe pumps, and the concentrations of solute molecules. Calculating the solute concentration requires information on the density of each solute molecule. However, since such data were not available, the density of each solute was assumed to be 1.0 g/cm^3, and theoretical concentrations were calculated using the molecular weights of the respective compounds.

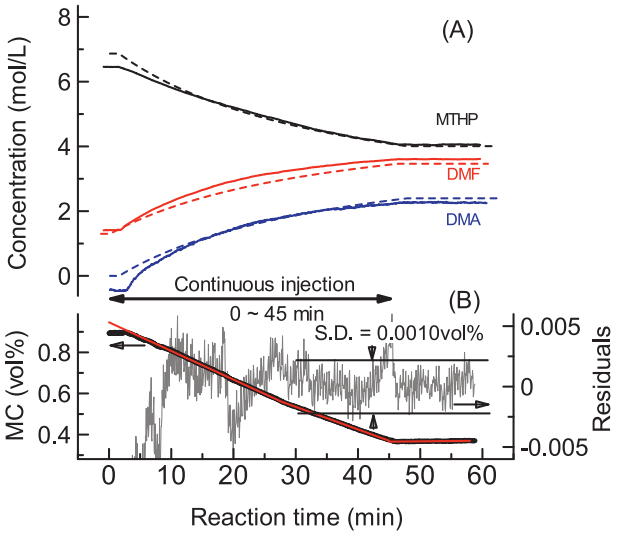

Figure 4 (A) Predicted solvent concentrations by PLSR (solid lines) and theoretical concentrations (dashed lines); (B) Prediction results for moisture content (MC) (black), fitted line of moisture variation (red), and the difference between predicted values and the fitted line (gray)

There was a delay in the concentration change due to retention in the flow path leading to the reaction vessel during the first 1.4 min after the start of liquid delivery. After that, the concentration of MTHP decreased, while those of DMF and DMA increased. The prediction results showed the same trend as theoretical values, with the predicted concentration of MTHP closely matching the theoretical value. The predicted values for DMF and DMA were also close to the theoretical values, and the concentrations after the cessation of liquid delivery (at 45 min) were also successfully predicted.

Figure 4(B) shows the prediction results for moisture content. A small amount of water was dissolved in the MTHP used for the STag starting material solution, and was present throughout the reaction. In this study, the moisture content was not measured directly, so the exact concentration could not be quantified. However, it is generally considered to be around 1.0 vol %, and the predicted value of 0.5 mol/L (0.9 vol %) was of a similar magnitude. In addition, the moisture content decreased linearly during the reaction as the reagent solutions were added, which is consistent with the fact that the added reagents contained no water. Assuming the initial concentration before the reaction was 0.50 mol/L (0.90 vol %, volume 18 mL), the final concentration at the cessation of liquid delivery was determined to be 0.22 mol/L (0.40 vol %) based on a final volume of 40.5 mL. This closely matched the predicted value of 0.20 mol/L (0.36 vol %).

Figure 4(B) also shows the result of linear fitting to the change in moisture content using the least squares method and calculating the difference between this fitted line and the predicted values. These differences represent the variation in the predicted values. The standard deviation (S.D.) was 0.0010 vol % (10 ppm). This value indicates high accuracy, even when compared with the results reported in the literature[11].

Prediction of Solvent Concentration in Flow Synthesis

After NIR spectral measurement began, delivery of the DIEA solution (dissolved in MTHP) and the COMU solution (dissolved in DMF) started at flow rates of 1.0 mL/min and 1.5 mL/min, respectively. Approximately 3 min later, delivery of Fmoc-Oic-OH (dissolved in DMA) and the STag starting material (dissolved in DMF/MTHP) started at flow rates of 3.0 mL/min and 12 mL/min, respectively. Figure 5 shows the NIR spectral measurements obtained during the flow synthesis. At 2.5 min after the start of measurement a strong absorption peak attributed to DMF was observed near 4980 cm^{-1}. At this point, the solvent mixture contained both MTHP and DMF, with DMF present in excess. At 5 min, the STag starting material, which was being delivered at a higher flow rate, merged into the stream and reached the flow cell, resulting in the observation of MTHP peaks (4700 cm^{-1} and 4780 cm^{-1}). From 7 min onward, mixing with Fmoc-Oic-OH and the coupling reaction began, and stable spectra were obtained.

Figure 6 shows the results of predicting each solvent concentration during flow synthesis using the spectra obtained and the regression model constructed by PLSR. Immediately after the start of measurement, the predictions were unstable due to the presence of residual solvents (water and isopropanol) and air bubbles in the flow synthesis system. However, DMF was detected at 2.5 min after the start of measurement. The prediction accuracy for all solvents was low. This was likely due to the DMF concentration exceeding the range of the regression model, as well as the flow rates not yet having stabilized. At 5 min, prediction accuracy remained low because the flow rate was still unstable, as it had been at 2.5 min. From 7 min onward, the flow rate stabilized, and the predicted solvent concentrations also became consistent. These

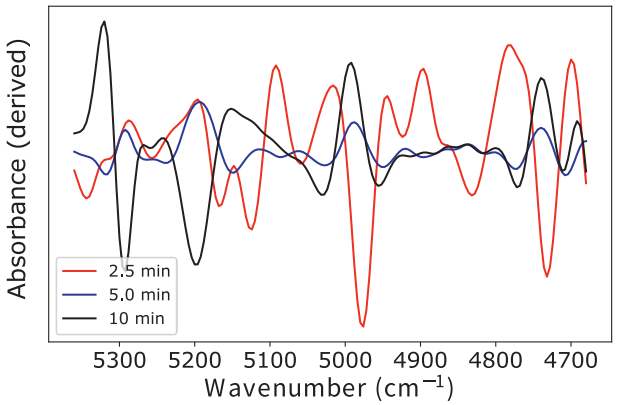

Figure 5 NIR spectra during flow synthesis: 2.5 min after measurement start (red), 5 min after measurement start (blue), 10 min after measurement start (black)

values closely approximated the theoretical concentrations, which were estimated based on the flow rates of each reagent solution and an assumed solute density of 1.0 g/cm^3.

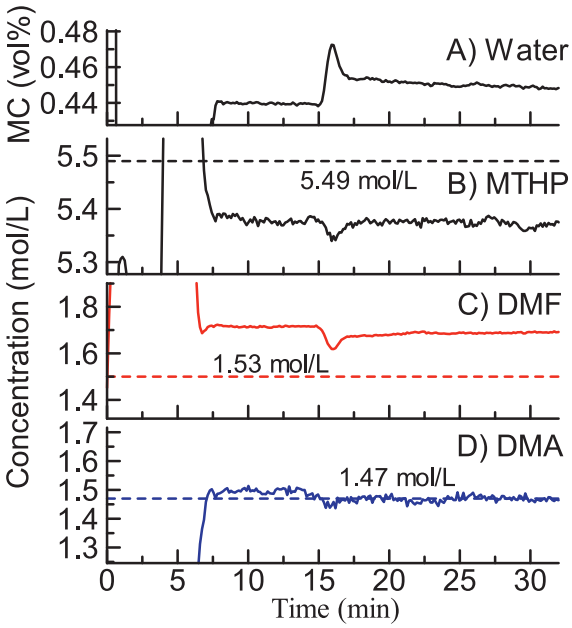

Figure 6 Predicted concentrations (mol/L) of solvents and moisture content (MC, vol%) during flow synthesis (solid lines) and theoretical concentrations (dashed lines)

Because a slight fluctuation in concentration was observed at 16 min after the start of measurement, the standard deviation (S.D.) of the measured values between 20 and 30 min was calculated. As shown in Table 3, the variation for all solvents was within 0.1 vol %. Flowmeter data confirmed that the flow rate and pressure for each solvent were stable, indicating that the solvent concentrations were steady and that the predicted values obtained by NIR were also stable.

The moisture content during flow synthesis was 0.45 vol %. At 16 min after the start of measurement, a slight increase in moisture was observed. This was likely due to the passage of micro-droplets of water contained in the STag

Table 3 Predicted solvent concentrations during the 20–30 min measurement interval

	Average		S.D.	
	mol/L	vol%	mol/L	vol%
MTHP	5.37	62.5	0.0063	0.073
DMF	1.69	13.0	0.0032	0.025
DMA	1.47	13.6	0.010	0.093
Water	0.25	0.45	0.00057	0.0010

starting material. An abnormal increase in moisture content can have a significant impact on the synthesis. It may lead to quality issues, such as reduced solubility of lipophilic compounds and inhibition of reactions, and can also result in blockage of the flow path. Therefore, monitoring is necessary. Based on the results in Figures 4 and 6, the standard deviation of the predicted moisture content was approximately 10 ppm. This indicates that even slight fluctuations in moisture content can be detected during flow synthesis, demonstrating the effectiveness of this method for monitoring synthesis processes.

CONCLUSION

In this study, a micro-flow cell was developed for real-time NIR measurement in batch and flow synthesis processes, and the prediction of solvent concentration and moisture content using PLSR was investigated. It was demonstrated that both solvent concentration and moisture content can be predicted with high accuracy. Solvent concentration serves as one indicator for confirming that a reaction is proceeding normally. For example, when an abnormality occurs in the reaction process, such as a decrease in product yield or an increase in impurities, solvent concentration can provide important information for identifying the cause.

It is known that moisture content can be quantified using NIR or IR spectra, and such methods have already been put into practical use. However, there are few examples of their application to synthesis processes. Water molecules have a low molecular weight, so even small amounts can affect reactions, necessitating high-precision detection. In the results of this study, the standard deviation of the predicted values was approximately 10 ppm, indicating a level of detection accuracy suitable for practical use.

In general, the primary purpose of real-time in-line spectroscopic analysis using NIR spectra and similar methods is to monitor raw materials and reaction products. In this study, it was demonstrated that solvent concentration and moisture content can also be predicted from the acquired spectra. This indicates that, depending on how it is applied, an in-line spectroscopic analysis system can provide a lot of information about the reaction, making it useful for understanding reaction mechanisms, examining reaction conditions, and detecting abnormalities. Moving forward, efforts will be made to expand its application to various areas in order to fully realize the potential of in-line spectroscopic analysis and enhance its value.

ACKNOWLEDGEMENT

This study was conducted in collaboration with PeptiStar, Inc.

REFERENCES

(1) Joshua Britton and Colin L. Raston, “Multi-step continuous-flow synthesis,” Chemical Society Reviews, Vol. 46, No. 5, 2017, pp. 1250-1271
(2) Shu Kobayashi, “Flow “Fine” Synthesis: High Yielding and Selective Organic Synthesis by Flow Methods,” Chemistry—An Asian Journal, Vol. 11, No. 4, 2015, pp. 425-436
(3) Naoto Sugisawa, Hiroyuki Nakamura, et al., “Recent Advances in Continuous-Flow Reactions Using Metal-Free Homogeneous Catalysts,” Catalysts, Vol. 10, No. 11, 2020, p. 1321
(4) Toshiyuki Kawabe, et al., ed. by Technical Information Institute, Scale-up and Continuous Operation of Chemical Processes, Technical Information Institute, Japan, 2019 (in Japanese)
(5) Koji Machida, Hideo Kawauchi, et al., “Development and Implementation of Innovative Technologies for Continuous Pharmaceutical Production,” Proceedings of the Summer Symposium of the Japanese Society for Process Chemistry, 2024, pp. 58-59 (in Japanese)
(6) Yusuke Hattori, “Raman and near-infrared spectroscopy for in-line sensors,” Analytical Sciences, Vol. 38, No. 11, 2022, pp. 1455-1456
(7) Shinya Yano, Toshihiro Mori, et al., “Silylated Tag-Assisted Peptide Synthesis: Continuous One-Pot Elongation for the Production of Difficult Peptides under Environmentally Friendly Conditions,” Molecules, Vol. 26, No. 12, 2021, p. 3497
(8) Daisuke Kubo, Yuma Otake, et al., “Development of a Continuous-Flow Manufacturing Device for Peptide Active Pharmaceutical Ingredients,” Yokogawa Technical Report English Edition, Vol. 66, No. 1, 2023, pp. 11-16
(9) V. Fornés, J. Chaussidon, “An interpretation of the evolution with temperature of the v_2+v_3 combination band in water,” The Journal of Chemical Physics, Vol. 68, No. 10, 1978, pp. 4667-4671
(10) Harvey Franklin Fisher, William C. McCabee, et al., “Near-infrared spectroscopic investigation of the effect of temperature on the structure of water,” The Journal of Physical Chemistry, Vol. 74, No. 25, 1970, pp. 4360-4369
(11) Salvador Garrigues, Máximo Gallignani, et al., “Flow-injection determination of water in organic solvents by near-infrared spectrometry,” Analytica Chimica Acta, Vol. 281, No. 2, 1993, pp. 259-264