

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Enhanced Copolymer Characterization for Poloxamers Using Gel Permeation Chromatography Combined with Artificial Neural Network

Journal:	<i>Analytical Chemistry</i>
Manuscript ID	ac-2022-02913m.R2
Manuscript Type:	Article
Date Submitted by the Author:	09-Feb-2023
Complete List of Authors:	Nagy , Tibor; University of Debrecen, Department of Applied Chemistry Roth, Gergo; University of Debrecen Benedek, Máté; University of Debrecen Kuki, Ákos; Debreceni Egyetem, Applied Chemistry Timári, István; Debreceni Egyetem, Department of Inorganic and Analytical Chemistry Zsuga, Miklos; Debreceni Egyetem, Applied Chemistry Kéki, Sándor; Debreceni Egyetem, Applied Chemistry

SCHOLARONE™
Manuscripts

Enhanced Copolymer Characterization for Poloxamers Using Gel Permeation Chromatography Combined with Artificial Neural Network

Tibor Nagy¹, Gergő Róth^{1,2}, Máté Benedek¹, Ákos Kuki¹, István Timári³, Miklós Zsuga¹,
Sándor Kéki^{1,*}

¹*Department of Applied Chemistry, Faculty of Science and Technology, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary*

²*Doctoral School of Chemistry, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary*

³*Department of Organic Chemistry, Faculty of Science and Technology, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary*

*E-mail: keki.sandor@science.unideb.hu.

Abstract

Gel permeation chromatography (GPC) is a generally applied method for the mass analysis of various polymers and copolymers, but it inherently fails to provide additional important information such as the composition of copolymers. However, we will show that GPC measurements using different solvents can yield not just the correct molecular weight, but the composition of the copolymer. Accordingly, artificial neural networks (ANNs) have been developed to process the data of GPC measurements and determine the molecular weight and the chemical composition of the copolymers. The target values of the ANNs were obtained by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) and nuclear magnetic resonance (NMR) spectroscopy. Our GPC – ANN method is demonstrated by the analysis of various poloxamers, i.e., polyethylene oxide (PEO) - polypropylene oxide (PPO) block copolymers. Two ANNs were constructed. The first one (*ANN_1*) works in a wider mass range (from 900 Dalton to 12500 Dalton), while the second one (*ANN_2*) produces more output values. *ANN_2* can thus predict seven characteristic copolymer parameters, namely two average molecular weights, the average weight fraction of the EO unit, and four average numbers of the repeat units. The correlation between the experimentally obtained outputs and the predicted ones is high, $r > 0.98$. The accuracy of the ANNs is very convincing, both ANNs predict the number-average molecular weight (M_n) with

an accuracy below 5%. Furthermore, this work is the first step for creating an open database and applications extending the use of the GPC–ANN method for the analysis of copolymers.

Keywords: gel permeation chromatography (GPC), copolymer characterization, machine learning, artificial neural network, mass spectrometry

Introduction

The structural and functional versatility of copolymers with unique properties offers the opportunity for their diverse applications in various fields. Therefore, the synthesis and characterization of these complex polymeric systems have recently attracted extensive research interest. Functional block copolymers with desired architecture can now be synthesized by controlled/“living” radical polymerization techniques such as atom transfer radical polymerization (ATRP)¹⁻³, stable free-radical polymerization (SFRP)⁴, and reversible addition-fragmentation chain transfer (RAFT) polymerization⁵⁻⁷. The development of advanced functional materials requires not only the emergence of novel synthetic methods but also further inventions in polymer/copolymer characterizations. Thus, to understand the structure-property relationships of copolymers and to tailor their functionalities and properties, comprehensive and accurate characterization of such complex chemical systems is of paramount importance⁸. However, due to their chemical diversity and structural complexity, the characterization of copolymers is very challenging. The copolymers are complex chemical systems that can be characterized by distributions of at least two different types of repeat units, their molecular weight, chemical composition, and chain architecture. Furthermore, the analysis of novel functional copolymers requires special evaluation methods rendering their detailed characterization labor-intensive and time-consuming⁹⁻¹³. Moreover, mass spectrometry and hyphenated techniques enable the detailed characterizations of such complex copolymer systems¹⁴⁻¹⁶, but the interpretation of the results (e.g. the mass spectra with a huge number of m/z peaks) necessity the use of special data mining techniques or computer software tools¹⁷⁻²¹. Other molecular-weight determination methods such as light scattering, viscosimetry, and nuclear magnetic resonance (NMR) spectroscopy can measure only the average, bulk properties of polymers.

Gel permeation chromatography (GPC), or size exclusion chromatography (SEC), is generally applied for the characterization of polymers. GPC is, however, a relative method meaning that

the molecular weight and molecular weight distribution can be obtained from the calibration of the instrument with known standards. Its speed, simplicity, and relatively low cost of operation make GPC the most widely used method for molecular weight distribution analysis of polymeric systems. However, GPC has limitations for the characterization of copolymers as it separates the molecules according to their hydrodynamic radius that depends not only on the molecular weight but also on the chemical composition of the copolymer, i.e., the molar ratio of the co-repeating units and their sequence in the macromolecule. Therefore, copolymer fractions with different molecular weights and different chemical compositions may co-elute^{22,23}. Although GPC with multiple detectors can resolve this problem, however, it requires an independent detector signal for each type of repeating unit. Furthermore, the hydrodynamic radius of the copolymers also depends on the solvent used for separation and the chemical composition of the copolymer. Therefore, to adequately characterize such copolymer systems, the solvent-polarity – molecular mass – composition – hydrodynamic volume relationship should be established, which is, however, not feasible. As an alternative approach, machine learning (ML), an artificial intelligence (AI) method, can be used. By ML, the hidden trends and relationships in data can be discovered for solving complex problems.²⁴ Machine learning methods can be divided into two broad categories: supervised learners and unsupervised learners. Supervised learning develops a predictive model based on both input and output data by using training data with known labels^{25, 26}. For example, a model is created performing a regression fit for predicting the average molecular weight of a polymer, where all the training data have a known average molecular weight. On the other hand, unsupervised learning methods analyze and cluster unlabeled data based only on inputs. Artificial neural network (ANN) is one of the most widely used methods in machine learning, which mimics the cognitive process of the human brain. ANNs model neurons of a biological nervous system as mathematical functions which are connected by weighted links²⁷. Artificial neural networks (ANN) are employed for diverse and complex tasks in the field of chemistry spanning from molecular simulation and analytical chemistry to prediction and design of chemical reactions, metabolic engineering and drug discovery, and so forth²⁶⁻³⁶. In this article, we report a new efficient GPC-ANN-based method for the characterization of copolymers. Artificial neural networks have been developed that process the data of three independent GPC measurements and determine the molecular weight and chemical composition of the copolymers.

Experimental Section

Chemicals

Copolymers and homopolymers produced by BASF were generous gifts (Ludwigshafen, Germany). Copolymers signed (A) were received in 2006, while the others were obtained in 2021. Further samples were purchased from Merck (Darmstadt, Germany) and Fluka (Buchs, Switzerland) details and the origin of the samples are summarized in the Supporting Information as Table S1. Methanol (MeOH), acetonitrile (can), dimethylformamide (DMF), tetrahydrofuran (THF), and dimethyl sulfoxide (DMSO) were received from VWR International (Leuven, Belgium). Water was purified by a Direct-Q water system (Millipore, Molsheim, France). In order to extend the number of samples, blends were prepared from the poloxamer analogues and flash chromatography was applied to fractionate copolymers. The suppliers of the samples are shown in the Supporting Information (Table S1).

Artificial Neural Network (ANN)

We developed an ANN model based on supervised learning which is the most common form of machine learning^{26,37,38}. The ANN modeling was performed using the Neural Network Toolbox of MATLAB 2020b (The MathWorks, Inc., Natick, MA, USA). Two feed-forward networks (*ANN_1* and *ANN_2*) with sigmoid hidden neurons and linear output neurons were constructed. The architecture of the ANNs will be detailed later. The networks were trained using the Bayesian regularization algorithm. The input data were randomly divided into training (90%) and test (10%) sets. The performance of the networks was evaluated by the mean square error (MSE) and Pearson's correlation coefficient (*r*) and was optimized by varying the number of hidden neurons (1-15). When the increase of hidden neurons did not yield better performance, the network with the smallest number of neurons at the maximum performance was chosen.

The evaluation of the ANN was based on the entire data set (the training plus test data). The accuracy (%) of the ANN prediction for the average molecular weights was calculated as the average of $abs(1 - predicted\ value / experimental\ target\ value) \times 100$ values for the entire data set. In order to avoid the division by zero, the accuracy of the prediction for the EO% and the average numbers of the repeat units were expressed as the average difference between the experimental and predicted values.

Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS)

The measurements were carried out by an Autoflex Speed MALDI-TOF MS instrument (Bruker Daltonik, Bremen, Germany) instrument. Both reflectron and linear modes were applied. In the reflectron mode, the ion source voltage 1, ion source voltage 2, reflector voltage 1, and reflector voltage 2 were 19 kV, 16.65 kV, 21 kV and 9.55 kV, respectively. In the linear mode the ion source voltage 1, and ion source voltage 2 were 19.5 kV and 18.3 kV. The mass spectrometer was equipped with a solid-phase laser (355 nm). External calibration was applied using polyethylene glycol and polypropylene glycol homopolymers. *Trans*-2-[3-(4-*tert*-Butylphenyl)-2-methyl-2-propenylidene]malononitrile (DCTB) matrix and sodium trifluoroacetate (NaTFA) ionizing agent dissolved in methanol were used for sample preparation. The concentrations of the matrix, samples, and ionizing agents were 20 mg/mL, 10 mg/mL, and 5 mg/mL, respectively, while the mixing ratio was 25:2:1. 0.5 L mixture was deposited on the surface of the MALDI plate and allowed to air-dry.

Gel permeation Chromatography

The gel permeation chromatographic measurements were carried out with a Waters Alliance e2695 separation module. For detection, a Waters 2414 refractive index detector was applied (Waters Corp., Milford, MA, USA). Three different eluents and methods were used. In the first set, DMF was the eluent (flow rate: 0.3 mL/min, T: 42°C, run: 28 min), and the applied columns were the following Shodex KF-603 and shodex KF-602.5 (150×6 mm, Showa Denko GmbH Munich, Germany). In the second set, THF,/DMSO/piperidine (94/5/1 V/V%) eluent was used (flow rate: 0.5 mL/min, T: 40°C, run: 35 min), columns: Styragel HR 0.5, 1, 2 and 4 (300×4.6 mm, Waters, Milford, MA, USA). In the third set, ACN:H₂O (60/40 V/V%) was employed as the eluent (flow rate: 0.48 mL/min, T: 40°C, run: 48 min, columns: two PolySep-GFC-P 3000 columns were coupled (length 300×7.8 mm, Agilent California, USA).

All GPC methods were calibrated by polyethylene glycol standards. Molecular weights (determined by MALDI-TOF MS) and the origin of the standards are shown in Supporting Information (Table S2.). Samples were dissolved in the eluents and filtered by a syringe filter (0.45 µm) prior to analysis with a concentration of approximately 10 mg/mL.

Nuclear magnetic resonance spectroscopy (NMR)

NMR experiments were performed on a Bruker Avance II 500 MHz spectrometer equipped with a 5 mm z-gradient TXI probe (Bruker Daltonik, Bremen, Germany). Quantitative 1D ^1H NMR spectra were obtained at 300 K with a single scan, a 90 degree ^1H excitation pulse, a spectral width of 14.97 ppm and 32768 total data points. An exponential window function with 0.3 Hz line broadening was applied to the time data before Fourier transform. Bruker TopSpin computer software (version 2.1 and 4.1) was used for processing, baseline correction, and integration of NMR spectra.

Results and Discussion

We demonstrate our new GPC – ANN method through the analysis of poly(ethylene oxide), poly(propylene oxide) homopolymers and poloxamers, i.e., poly(ethylene oxide) – block - poly(propylene oxide) – block - poly(ethylene oxide) (b-PEO–b-PPO–b-PEO) copolymers and the reversed b-PPO–b-PEO–b-PPO structures. These amphiphilic block copolymers are extensively used as nonionic surfactants in industrial and domestic applications ranging from pharmaceuticals to cosmetics, and so forth^{39–41}. In the first set of our experiments, we analyzed 61 polymers (12 homopolymers and 49 poloxamers) with various number-average molecular weights (M_n) ranging from 900 g/mol to 12500 g/mol and with different ethylene oxide (EO) content, i.e., the different weight fractions of EO unit (EO%). The M_n values were determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The average weight fractions of EO were determined by nuclear magnetic resonance (NMR) spectroscopy. The M_n and EO% values were used for the training and testing of the ANNs, as will be detailed later. Figure 1 shows the measured M_n and EO% values for 16 representative copolymers together with the manufacturers' specifications arranged in a 2D plot. As can be seen in Fig. 1, for some copolymers, the measured M_n and/or EO% values (blue dots) differ from the expected ones (red dots). Additional examples of this deviation can be found in our previous paper²¹, where we also demonstrated that small variations in the copolymer composition significantly affect the copolymer properties, for example, the aggregation behavior of the amphiphilic block copolymers. It emphasizes the importance of reliable copolymer characterization methods, even if their data sheets are available.

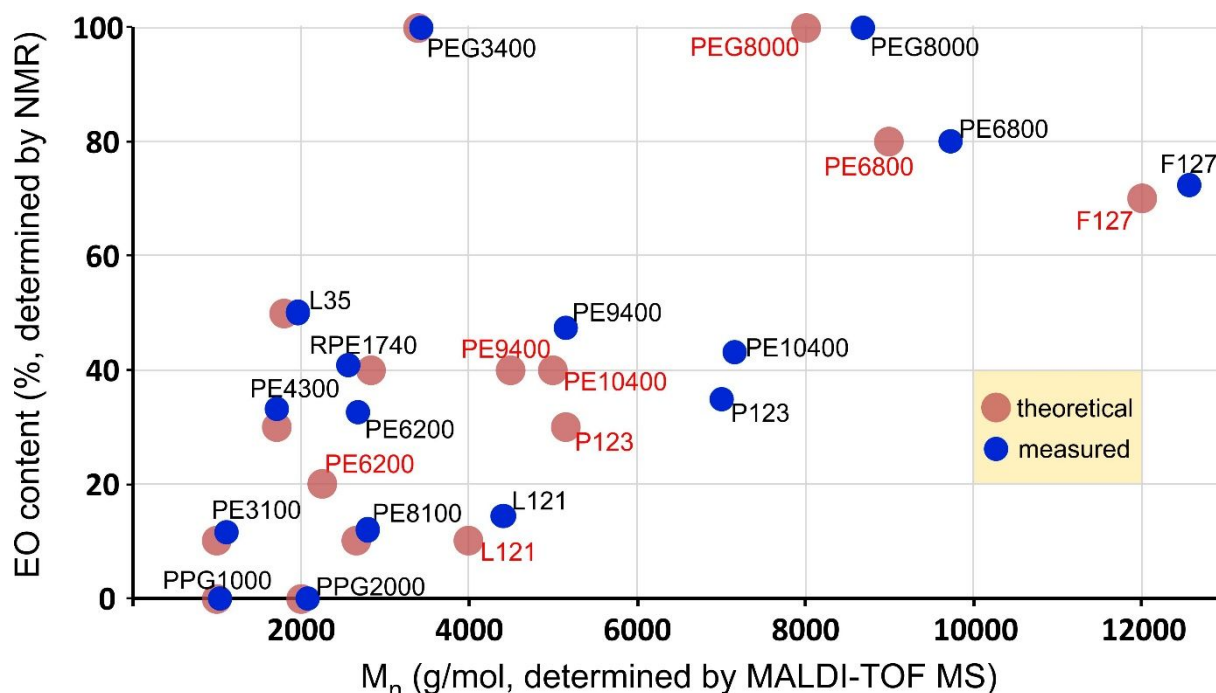


Figure 1. The number average molecular weights (M_n , measured by MALDI-TOF MS) versus EO wt% (determined by NMR spectroscopy) of 16 selected poloxamers.

Figure 2 shows the molecular weight distributions of the reverse triblock Pluronic RPE3110 with approximately 3500 g/mol number-average molar mass and 10 wt % EO content recorded by GPC using three different solvents. The remarkable shifts of the distributions in Figure 2 clearly demonstrate that the hydrodynamic radius of the copolymers strongly depends on the polarity of the solvent used. It must be noted that in some cases bimodal distribution was seen on the GPC traces. Impurities in poloxamers were reported, however, the relatively low percent of the impurity results in only small differences between the bulk and fractions of the main components⁴². As an example, the GPC trace of PE9400 copolymer is shown in Supporting Information as Figure S1. Moreover, the effect of solvent polarity on the hydrodynamic radius and thus, on the GPC results are influenced by the hydrodynamic volume and the chemical composition of the copolymer, as it is demonstrated in Figure 3 which shows the differences between the M_n values measured by MALDI-TOF MS and GPC (with PEG calibration). Consequently, the dependence of the GPC traces on the solvent polarity carries information on the molecular weight distribution and the composition (e.g. block length distribution) of the copolymers. Our idea was to construct ANN models that receive inputs from the GPC and deliver the main copolymer characteristics as outputs. Excellent reproducibility of the GPC measurements was obtained for the 31R1, RPE2525, PE6800, and L121 copolymers. The standard deviation and RSD% are shown in the Supporting information (Table S3). Thus for further analysis single injection was used.

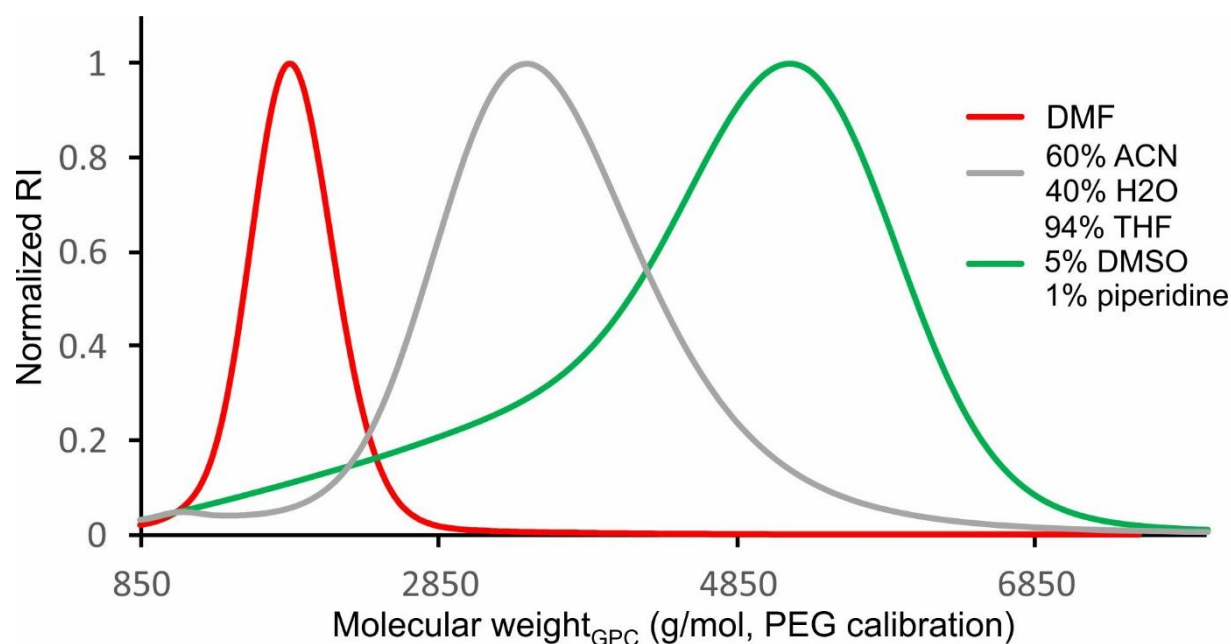


Figure 2. The GPC chromatogram of the RPE3110 poloxamer measured with different eluents. The molecular weight was calculated based on the retention times by applying PEG standards for calibration.

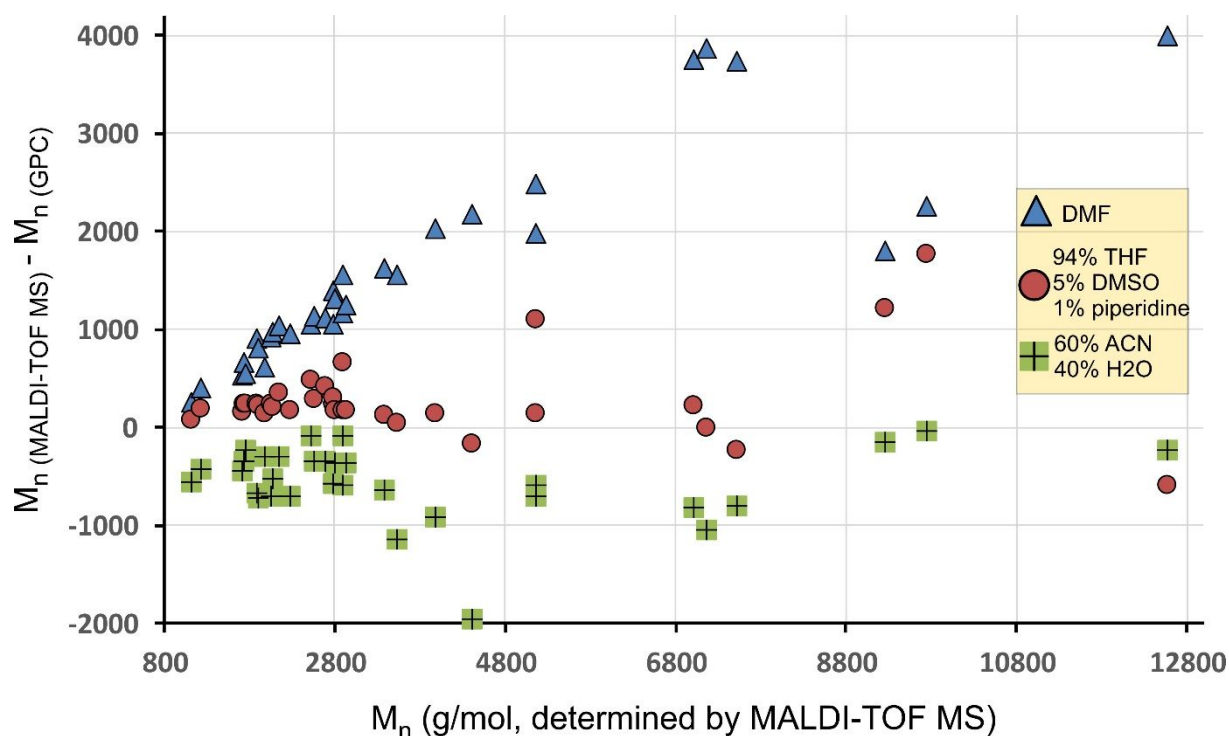


Figure 3. The differences between the number-average molecular weights measured by MALDI-TOF MS and the number-average molecular weights identified by GPC with different eluent and PEG standards for calibration.

First, a two-layer feed-forward ANN (ANN_I) with three sigmoid hidden neurons and two linear output neurons was constructed (Figure 4). The network has three input variables altogether which are the M_p values from the GPC chromatograms. The visualized summary of the measurements and dataset for ANN_I is shown in the Supporting Information (Figure S2). The outputs of the ANN are the predicted number-average molecular weight (M_n) and the ethylene oxide content (i.e., the average weight fraction of EO, EO%). The number of hidden neurons was determined as detailed in the Experimental Section. The targets of the ANN were the M_n and EO% values determined by MALDI-TOF MS and NMR spectroscopy, respectively. The polymer quantities, measured and predicted by ANN_I , are summarized in the Supporting Information (Table S4).

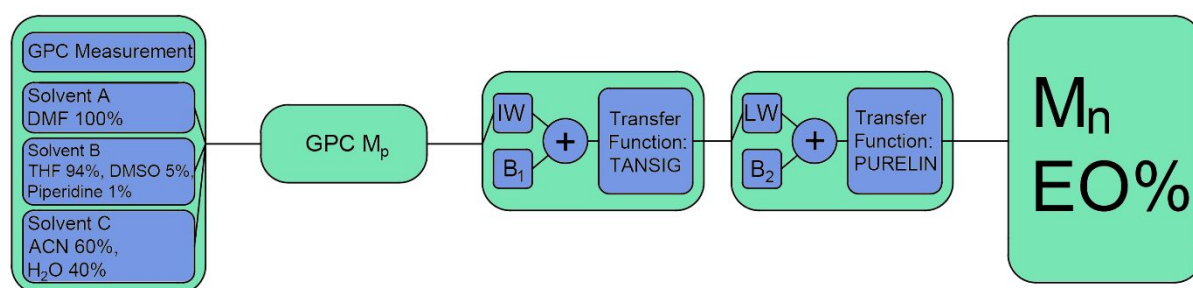


Figure 4. ANN_I with a 3-3-2 architecture used for the M_n and EO% determination.

The neural network was trained, validated, and tested using the GPC, MALDI-TOF MS, and NMR data of 61 polymers. The MATLAB regression plots displaying the network outputs concerning targets for training, testing, and all data sets are shown in the Supporting Information (Figure S3). Figure 5a and 5b reveal the parity plots comparing the experimental number-average molecular weight, and the ethylene oxide content, respectively, to the corresponding, predicted results for all data sets.

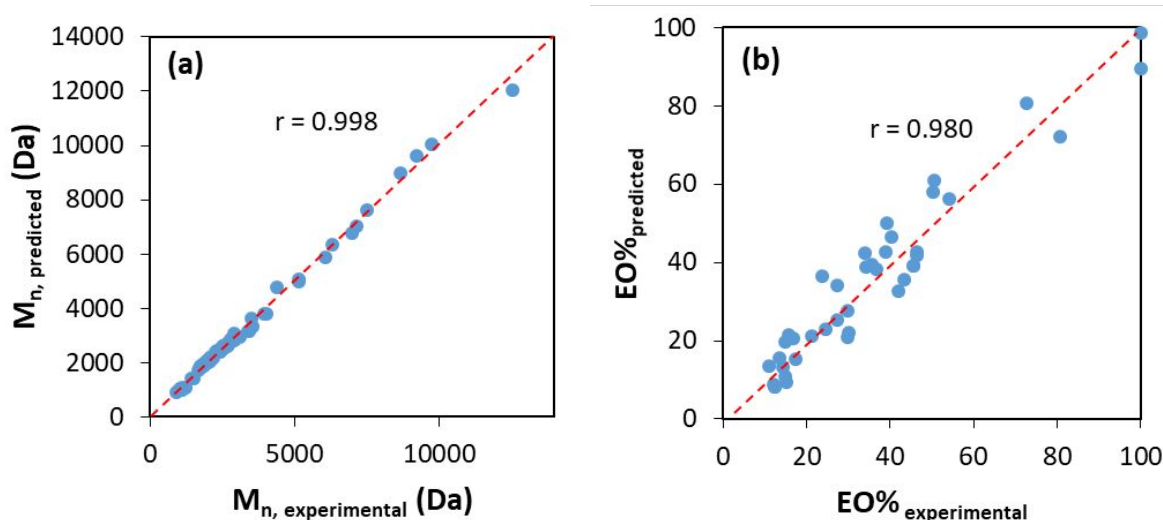


Figure 5. Parity plot (a) between the M_n experimentally obtained by MALDI-TOF MS and the M_n predicted by ANN_1, (b) between the EO% experimentally obtained by NMR and the EO% predicted by ANN_1.

As seen in Figure 5, the correlation between the experimental and the predicted values is very good, especially for the M_n data, the Pearson's correlation coefficients are $r = 0.998$ and $r = 0.980$ for the M_n and EO% values, respectively. Moreover, the predicted M_n values lie on the diagonal line showing that ANN_1 accurately predicts the number-average molecular weights (Figure 5a). The EO% values fall fairly close to the diagonal, indicating an acceptable accuracy. The accuracy of the ANN prediction can be quantified, as detailed in the Experimental Section. Accuracy of 3.6% and 5.3 EO% were obtained for the M_n and EO% prediction, respectively (the latter is expressed as the average difference between the experimental and predicted values). The paired T-tests show no significant differences between the experimental and predicted M_n and EO% values ($p = 0.84$ and $p = 0.95$, respectively).

In the second series of experiments, 50 polymers (9 homopolymers and 41 poloxamers) with M_n up to 5000 g/mol were characterized. The lower mass range allows us the comprehensive characterization of the copolymer composition by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) combined with our recently invented data processing algorithm developed for the compositional assignment and the estimation of the relative abundance of all molecular species present in the copolymer mass spectra²¹. We also determined, in addition to the usual molecular weight averages M_n and M_w , the number-average number of EO and propylene oxide (PO) repeat units (n_n^{EO} , n_n^{PO}) and the weight-average number of EO and PO units (n_w^{EO} , n_w^{PO}). The average weight fractions of EO (EO%) were also

determined by $^1\text{H-NMR}$. It is to be noted that detailed evaluation of the copolymer mass spectra is not a straightforward task due to, for example, the presence of a huge number of m/z peaks, some of them overlapped or having isobaric composition. This was one of our motives to create ANNs, which make GPC capable of comprehensive characterization of copolymers. The n_n^{EO} , n_n^{PO} , n_w^{EO} , and n_w^{PO} , values from MS characterization and the M_p values from GPC measurements were used for the training and testing of *ANN_2*. The M_n , M_w and EO% values are easily derived from the training and testing data set.

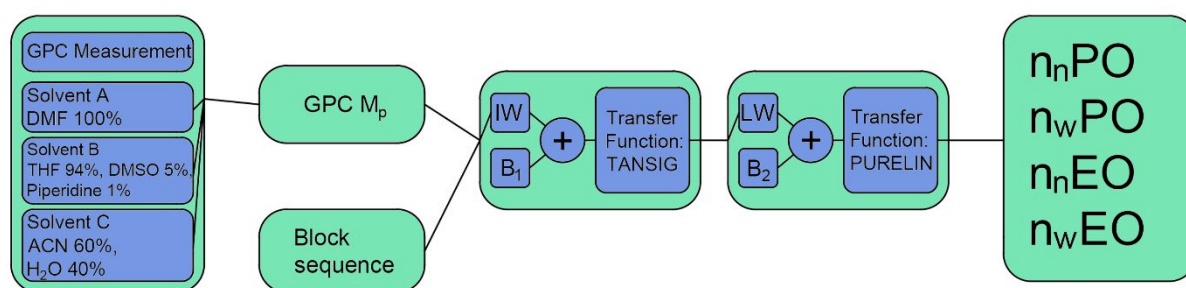


Figure 6. *ANN_2* with a 4-5-7 architecture used for the detailed characterization of the copolymers.

Figure 6 shows the architecture of the artificial neural network *ANN_2*. The two-layer feed-forward ANN has five sigmoid hidden neurons and four linear output neurons. *ANN_2* has four input variables, the M_p values from GPC measurement and the flag “Block sequence”. The outputs of *ANN_2* are the predicted n_n^{EO} , n_n^{PO} , n_w^{EO} , and n_w^{PO} , values. The number of hidden neurons was determined as detailed in the Experimental Section. The targets of *ANN_2* were the n_n^{EO} , n_n^{PO} , n_w^{EO} , and, n_w^{PO} results of the detailed MALDI-TOF MS characterization. *ANN_2* was trained, validated, and tested using the GPC, and MALDI-TOF MS data of 50 (co)polymers in the mass range from 900 g/mol to 5000 g/mol. The target variables (and additional copolymer quantities) of *ANN_2* are summarized in the the Supporting Information (Table S5). The MATLAB regression plots displaying the network outputs concerning targets for training, testing, and all data sets are shown in the Supporting Information (Figure S4). Figures 7a-7d show the parity plots comparing the experimental target values to the corresponding predicted results of *ANN_2* for all data sets.

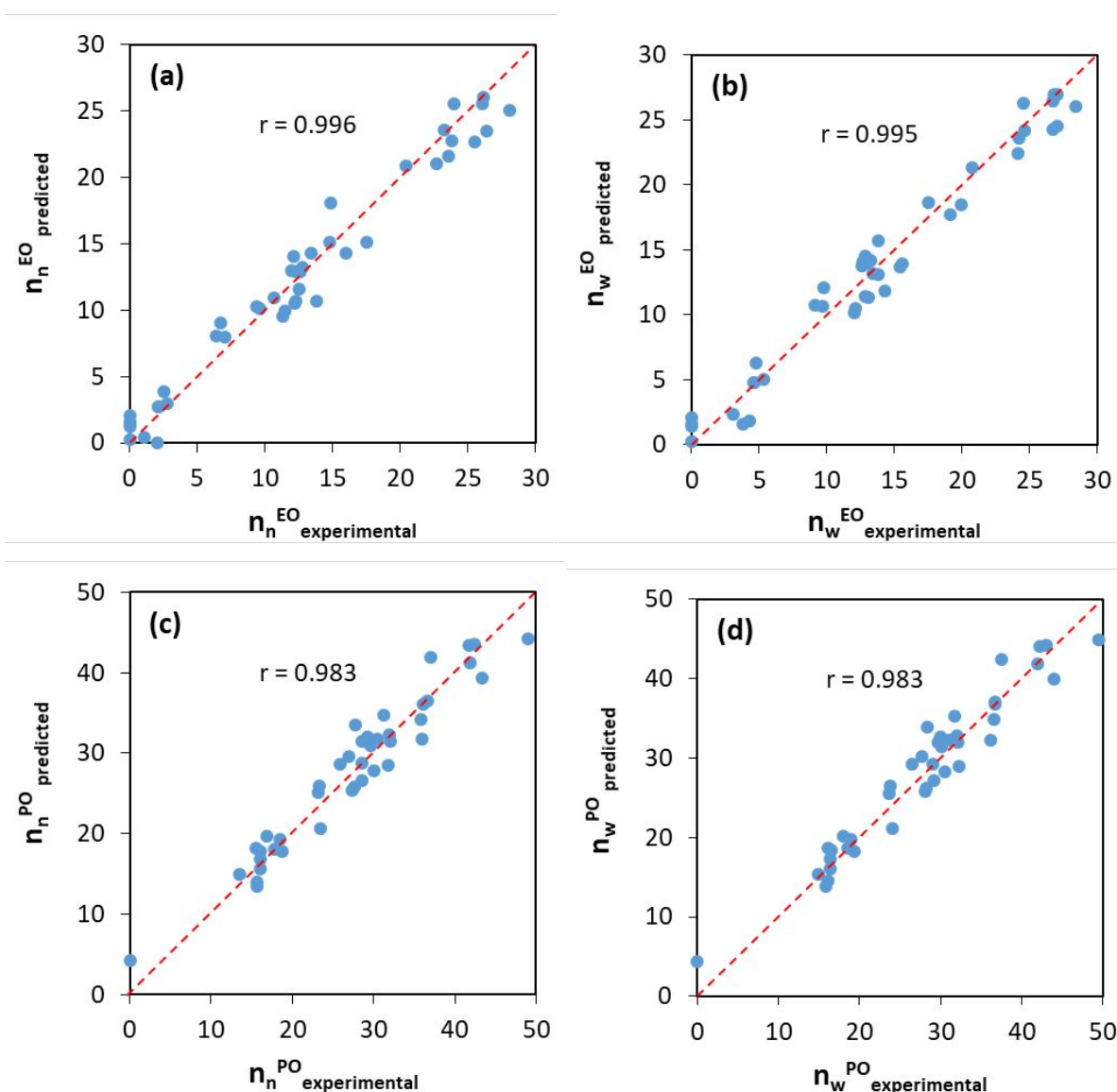


Figure 7. Parity plots between the experimentally obtained (a) n_n^{EO} , (b) n_w^{EO} , (c) n_n^{PO} , (d) n_w^{PO} values and the corresponding predicted values of *ANN_2*.

As seen in Figure 7, the correlations between the experimental and predicted values are excellent with Pearson's correlation coefficients greater than $r = 0.98$. It is also clear from Figure 7, that the accuracy of *ANN_2* is good because the predicted values lie close to the diagonal line. The validation data are summarized in Table 1.

Table 1. Validation data of *ANN_2*: Pearson's coefficient r of the correlation between the experimental and predicted values; the accuracy of the prediction (see Experimental Section); p -value of the paired T-test between the experimental and predicted values.

	M_n	M_w	EO%	n_n^{EO}	n_w^{EO}	n_n^{PO}	n_w^{PO}
r	0.988	0.987	0.990	0.996	0.995	0.983	0.983
Accuracy	4.9*	4.1*	2.3	1.6	1.6	2.2	2.2
p	0.95	0.83	0.51	0.09	0.08	0.34	0.35

* Percentage, as detailed in the *Experimental*.

As seen in Table 1, the quantified validation data also show a very high-quality prediction by *ANN_2*, the accuracies are below 5% for the average molecular weights; the paired T-tests show no significant differences between the experimental and predicted values.

In our previous paper, we introduced a novel copolymer quantity, namely the polydispersity ratio (*PDR*) of the comonomers²¹. *PDR* is the ratio of the polydispersity indices of the repeat units. In the case of EO/PO block copolymers, *PDR* is calculated as $PDR = PDI_{EO} / PDI_{PO}$, where $PDI_{EO} = n_w^{EO} / n_n^{EO}$ and $PDI_{PO} = n_w^{PO} / n_n^{PO}$. As we have shown earlier²¹, *PDR* can predict the presence of diblock impurities in a triblock copolymer, for example, $PDR > 1$ indicates the presence of diblocks in the PEO-PPO-PEO poloxamers. Herein, we extend the scope of *PDR*, because our ANN-based method is capable of the determination of copolymer quantities n_n^{EO} , n_n^{PO} , n_w^{EO} , and n_w^{PO} simply from the GPC results. Figure 8 demonstrates the correlation between the percentages of the secondary OH ends measured by NMR and the *PDR* values determined by our GPC-ANN method, both indicate the presence of PEO-PPO diblocks in the PEO-PPO-PEO copolymer sample. It can also be surmised from Figure 8 that the presence of approximately 100% secondary OH group in the PPO-PEO-PPO copolymers is in line with $PDR = 1$, i.e., suggesting also the absence of PPO-PEO diblocks.

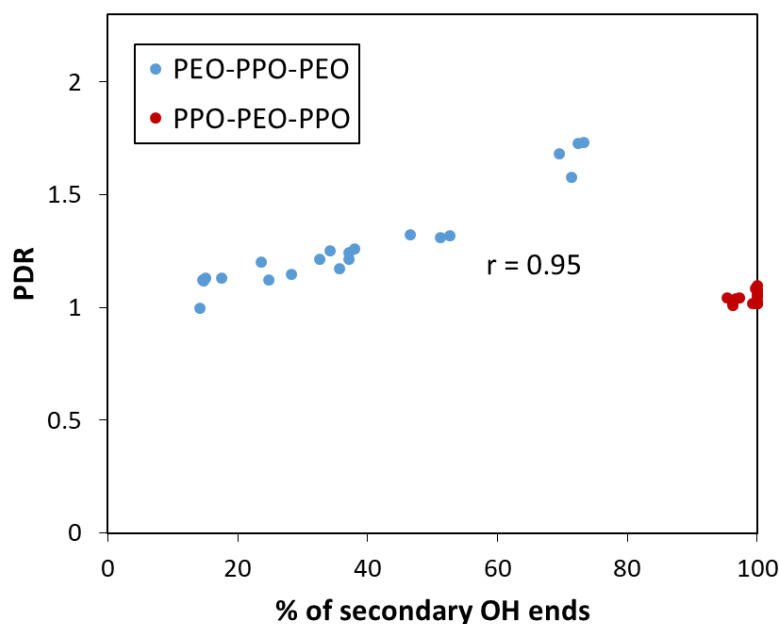


Figure 8. Correlation of the PDR values and the presence of secondary OH end-groups. The PDR values were determined by the GPC-*ANN_2* method while the secondary OH end-groups were measured by ^1H -NMR spectroscopy. r stands for Pearson's correlation coefficient.

Conclusions

In this paper, we have proposed a new gel permeation chromatography–artificial neural network (GPC-ANN) method for the detailed characterization of copolymers. The key was running multiple GPC measurements using eluents with different polarities. The ANN was trained and validated using the accurate chemical composition of the copolymers determined by MALDI-TOF MS and ^1H -NMR. We constructed two ANNs, the first (*ANN_1*) is capable of the characterization of copolymers with a wider mass range, the second (*ANN_2*) is limited up to the masses of 5000 g/mol, but this latter provides more sophisticated copolymer characterization. Both ANNs correlate at a very high level comparing the experimental target values to the corresponding predicted results (the r coefficients are at least 0.98 for most of the output variables). The accuracy of the prediction is remarkably high, e.g. for the M_n it is 3.6% and 4.9% for *ANN_1* and *ANN_2*, respectively. The main benefit of our work is that the database containing our GPC, MALDI-TOF MS, and ^1H -NMR data can be accessed and extended with additional experiments and thus the mass range and accuracy of our ANNs can further be improved. All the measured data and excel documents are available making the direct

application of the developed ANNs possible. It means that detailed characterization of a poloxamer can be carried out by the use of three GPC methods combined with our ANNs. Furthermore, the proposed method can easily be adopted for the characterization of other copolymers with various structures. This work is the first step for creating an open database and applications extending the use of the GPC–ANN method for the analysis of copolymers.

Associated content

Supporting Information

Implemented_ANN_1.xlsx The implemented ANN_1 into an excel sheet, any measured data can be converted to predicted polymer quantities.

Implemented_ANN_2.xlsx The implemented ANN_2 into an excel sheet, any measured data can be converted to predicted polymer quantities.

Supporting_Information.docx

Table S1 The list and type of copolymers involved in the analysis, Table S2 The PEG standards used for calibration for the gel permeation chromatography, Table S3 The reproducibility M_n , M_w and MP values of four copolymer samples, determined from 7 injections (applying DMF as eluent). Table S4 The input parameters of the neural network ANN_1 for the identification of molecular weight and chemical composition. MALDI-MS and ^1H -NMR data are the measured data. Table S5 The parameters of the neural network ANN_2 for the detailed characterization of the copolymers, Figure S1 The GPC chromatogram of the PE9400 copolymer measured with THF/DMSO/piperidine (94/5/1 V/V%) eluent. Figure S2 The summary of measurements and their results and the classification of the data which was used as input and output data for the supervised ANNs. Figure S3 MATLAB regression plots displaying ANN_1 outputs with respect to targets for training, testing and all data sets. Figure S4 MATLAB regression plots displaying ANN_2 outputs with respect to targets for training, testing and all data sets.

Author Information

Corresponding Author

*E-mail: keki.sandor@science.unideb.hu. Fax: +36 52 518662

ORCID

Sándor Kéki: 0000-0002-5274-6117

Tibor Nagy: 0000-0001-8568-914X

Acknowledgments

The work was supported by the GINOP-2.3.3-15-2016-00021 projects and by Thematic Excellence Program (TKP2020-NKA-04) of the Ministry for Innovation and Technology in Hungary. Furthermore, this paper was also supported by Grant No. FK-132385 from National Research, Development and Innovation Office (NKFI).

T.N. and I.T. acknowledge the support of the National Research, Development and Innovation Office of Hungary (PD 135034 (I. T.)), the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/00212/20/7 (T.N.), BO/00372/20/7 (I.T.)), ÚNKP-22-05-DE-426 (T.N.), ÚNKP-22-5-DE-424 (I.T.) and the ÚNKP-21-2-II-DE-227 (M.B.) New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund. Thanks to the BASF (Germany) for the numerous samples obtained as gifts, and thanks for the financial support provided by the Mol Group, Hungary.

References

1. Treat, N. J.; Sprafke, H.; Kramer, J. W.; Clark, P. G.; Barton, B. E.; Read de Alaniz, J.; Fors, B. P.; Hawker, C. J., Metal-Free Atom Transfer Radical Polymerization. *J. Am. Chem. Soc.* **2014**, *136* (45), 16096-16101.
2. Herberg, A.; Yu, X.; Kuckling, D., End Group Stability of Atom Transfer Radical Polymerization (ATRP)-Synthesized Poly(N-isopropylacrylamide): Perspectives for Diblock Copolymer Synthesis. *Polymers* **2019**, *11* (4), 678.
3. Matyjaszewski, K., Advanced Materials by Atom Transfer Radical Polymerization. *Adv. Mater.* **2018**, *30* (23), 1706441.
4. Hansen, K.-A.; Blinco, J. P., Nitroxide radical polymers – a versatile material class for high-tech applications. *Polymer Chemistry* **2018**, *9* (13), 1479-1516.
5. Perrier, S., 50th Anniversary Perspective: RAFT Polymerization—A User Guide. *Macromolecules* **2017**, *50* (19), 7433-7447.
6. Gody, G.; Barbey, R.; Danial, M.; Perrier, S., Ultrafast RAFT polymerization: multiblock copolymers within minutes. *Polymer Chemistry* **2015**, *6* (9), 1502-1511.
7. Jo, Y. S.; van der Vlies, A. J.; Gantz, J.; Antonijevic, S.; Demurtas, D.; Velluto, D.; Hubbell, J. A., RAFT Homo- and Copolymerization of N-Acryloyl-morpholine, Piperidine, and Azocane and Their Self-Assembled Structures. *Macromolecules* **2008**, *41* (4), 1140-1150.
8. Zhang, C.; Bates, M. W.; Geng, Z.; Levi, A. E.; Vigil, D.; Barbon, S. M.; Loman, T.; Delaney, K. T.; Fredrickson, G. H.; Bates, C. M.; Whittaker, A. K.; Hawker, C. J., Rapid Generation of Block Copolymer Libraries Using Automated Chromatographic Separation. *J. Am. Chem. Soc.* **2020**, *142* (21), 9843-9849.
9. Niezen, L. E.; Staal, B. B. P.; Lang, C.; Pirok, B. W. J.; Schoenmakers, P. J., Thermal modulation to enhance two-dimensional liquid chromatography separations of polymers. *J. Chromatogr. A* **2021**, *1653*, 462429.

10. Lee, S.; Choi, H.; Chang, T.; Staal, B., Two-Dimensional Liquid Chromatography Analysis of Polystyrene/Polybutadiene Block Copolymers. *Anal. Chem.* **2018**, *90* (10), 6259-6266.
11. Erlandsson, B.; Wittgren, B.; Brinkmalm, G., Development of a size exclusion chromatography method for the determination of molar mass for poloxamers. *J. Pharm. Biomed. Anal.* **2003**, *31* (5), 845-858.
12. Molenaar, S. R. A.; van de Put, B.; Desport, J. S.; Samanipour, S.; Peters, R. A. H.; Pirok, B. W. J., Automated Feature Mining for Two-Dimensional Liquid Chromatography Applied to Polymers Enabled by Mass Remainder Analysis. *Anal. Chem.* **2022**.
13. Meunier, D. M.; Wade, J. H.; Janco, M.; Cong, R.; Gao, W.; Li, Y.; Mekap, D.; Wang, G., Recent Advances in Separation-Based Techniques for Synthetic Polymer Characterization. *Anal. Chem.* **2021**, *93* (1), 273-294.
14. Crotty, S.; Gerişlioğlu, S.; Endres, K. J.; Wesdemiotis, C.; Schubert, U. S., Polymer architectures via mass spectrometry and hyphenated techniques: A review. *Anal. Chim. Acta* **2016**, *932*, 1-21.
15. Peacock, P. M.; McEwen, C. N., Mass Spectrometry of Synthetic Polymers. *Anal. Chem.* **2004**, *76* (12), 3417-3428.
16. Gruendling, T.; Weidner, S.; Falkenhagen, J.; Barner-Kowollik, C., Mass spectrometry in polymer chemistry: a state-of-the-art up-date. *Polymer Chemistry* **2010**, *1* (5), 599-617.
17. Nagy, T.; Kuki, Á.; Zsuga, M.; Kéki, S., Mass-Remainder Analysis (MARA): a New Data Mining Tool for Copolymer Characterization. *Anal. Chem.* **2018**, *90* (6), 3892-3897.
18. Fouquet, T.; Sato, H., Extension of the Kendrick Mass Defect Analysis of Homopolymers to Low Resolution and High Mass Range Mass Spectra Using Fractional Base Units. *Anal. Chem.* **2017**, *89* (5), 2682-2686.
19. Nagy, T.; Kuki, Á.; Hashimov, M.; Zsuga, M.; Kéki, S., Multistep Mass-Remainder Analysis and its Application in Copolymer Blends. *Macromolecules* **2020**, *53* (4), 1199-1204.
20. Engler, M. S.; Crotty, S.; Barthel, M. J.; Pietsch, C.; Knop, K.; Schubert, U. S.; Böcker, S., COCONUT—An Efficient Tool for Estimating Copolymer Compositions from Mass Spectra. *Anal. Chem.* **2015**, *87* (10), 5223-5231.
21. Róth, G.; Nagy, T.; Kuki, Á.; Hashimov, M.; Vonza, Z.; Timári, I.; Zsuga, M.; Kéki, S., Polydispersity Ratio and Its Application for the Characterization of Poloxamers. *Macromolecules* **2021**, *54* (21), 9984-9991.
22. Sebastian, J. M.; Register, R. A., Block copolymer molecular weight determination via gel permeation chromatography: Choosing a combining rule. *J. Appl. Polym. Sci.* **2001**, *82* (8), 2056-2069.
23. Kilz, P.; Held, D. M. In *Practical Copolymer Analysis Using GPC/SEC and related techniques*, 2017.
24. Moon, J.; Gbadago, D. Q.; Hwang, G.; Lee, D.; Hwang, S., Software platform for high-fidelity-data-based artificial neural network modeling and process optimization in chemical engineering. *Comput. Chem. Eng.* **2022**, *158*, 107637.
25. Ippolito, M.; Ferguson, J.; Jenson, F., Improving facies prediction by combining supervised and unsupervised learning methods. *Journal of Petroleum Science and Engineering* **2021**, *200*, 108300.
26. Ayres, L. B.; Gomez, F. J. V.; Linton, J. R.; Silva, M. F.; Garcia, C. D., Taking the leap between analytical chemistry and artificial intelligence: A tutorial review. *Anal. Chim. Acta* **2021**, *1161*, 338403.
27. Walczak, S.; Cerpa, N., Artificial Neural Networks. In *Encyclopedia of Physical Science and Technology (Third Edition)*, Meyers, R. A., Ed. Academic Press: New York, 2003; pp 631-645.
28. Gasteiger, J., Chemistry in Times of Artificial Intelligence. *Chemphyschem* **2020**, *21* (20), 2233-2242.
29. Schütt, K. T.; Chmiela, S.; von Lilienfeld, O. A.; Tkatchenko, A.; Tsuda, K.; Müller, K. R., *Machine Learning Meets Quantum Physics*. Springer International Publishing: 2020.
30. von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A., Exploring chemical compound space with quantum-based machine learning. *Nature Reviews Chemistry* **2020**, *4* (7), 347-358.
31. Behler, J., First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (42), 12828-12840.

32. Meuwly, M., Machine Learning for Chemical Reactions. *Chem. Rev.* **2021**, *121* (16), 10218-10239.
33. Packwood, D.; Nguyen, L. T. H.; Cesana, P.; Zhang, G.; Staykov, A.; Fukumoto, Y.; Nguyen, D. H., Machine Learning in Materials Chemistry: An Invitation. *Machine Learning with Applications* **2022**, *8*, 100265.
34. Zhou, Q.; Lu, S.; Wu, Y.; Wang, J., Property-Oriented Material Design Based on a Data-Driven Machine Learning Technique. *The Journal of Physical Chemistry Letters* **2020**, *11* (10), 3920-3927.
35. Kakkar, S.; Kwapinski, W.; Howard, C. A.; Kumar, K. V., Deep neural networks in chemical engineering classrooms to accurately model adsorption equilibrium data. *Education for Chemical Engineers* **2021**, *36*, 115-127.
36. Ekins, S.; Puhl, A. C.; Zorn, K. M.; Lane, T. R.; Russo, D. P.; Klein, J. J.; Hickey, A. J.; Clark, A. M., Exploiting machine learning for end-to-end drug discovery and development. *Nature Materials* **2019**, *18* (5), 435-441.
37. LeCun, Y.; Bengio, Y.; Hinton, G., Deep learning. *Nature* **2015**, *521* (7553), 436-444.
38. Cencer, M. M.; Moore, J. S.; Assary, R. S., Machine learning for polymeric materials: an introduction. *Polym. Int.* **2022**, *71* (5), 537-542.
39. Liu, T.; Wu, C.; Xie, Y.; Liang, D.; Zhou, S.; Nace, V. M.; Chu, B., Amphiphilic Polyoxyalkylene Triblock Copolymers: Self-Assembly, Phase Behaviors, and New Applications. In *Associative Polymers in Aqueous Media*, American Chemical Society: 2000; Vol. 765, pp 2-20.
40. Yokoyama, M., Block copolymers as drug carriers. *Crit. Rev. Ther. Drug Carrier Syst.* **1992**, *9* (3-4), 213-248.
41. Edens, M. W., Applications of Polyoxyalkylene Block Copolymer Surfactants In *Nonionic Surfactants Polyoxyalkylene Block Copolymers*, Nace, V., Ed. Taylor & Francis: 1996; p 26.
42. Batsberg, W.; Ndoni, S.; Trandum, C.; Hvidt, S., Effects of Poloxamer Inhomogeneities on Micellization in Water. *Macromolecules* **2004**, *37* (8), 2965-2971.

For Table of Content only

