

ORIGINAL ARTICLE

Multivariate analysis of ^{13}C NMR spectra of branched copolymers prepared by initiator-fragment incorporation radical copolymerization of ethylene glycol dimethacrylate and *tert*-butyl methacrylate

Tomohiro Hirano¹, Ryota Kamiike¹, Yuchin Hsu¹, Hikaru Momose^{1,2} and Koichi Ute¹

In this paper, we report chemometric approach for structural analysis of branched copolymers. To evaluate chemical compositions and degree of branching (DB) values in branched copolymers, multivariate analyses, such as principal component analysis (PCA) and partial least-squares (PLS) regression, were applied to the ^{13}C nuclear magnetic resonance (NMR) spectra of the carbonyl carbons of the copolymers prepared by initiator-fragment incorporation radical copolymerization of ethylene glycol dimethacrylate (EGDMA) and *tert*-butyl methacrylate (TBMA) with dimethyl 2,2'-azobisisobutyrate (MAIB). PCA successfully extracted information on monomeric units, such as EGDMA units, TBMA units and MAIB fragments, the last of which were incorporated via initiation and primary radical termination. The chemical compositions and DB values of the copolymers were predicted by PLS regression. Proper selection of a training set was found to be important for the prediction: the training set has to contain branched copolymers along with poly(EGDMA) and poly(TBMA). PLS regression using the appropriate training set allowed us to predict quantitatively the chemical compositions and DB values, without any assignments of the individual peaks.

Polymer Journal (2016) 48, 793–800; doi:10.1038/pj.2016.16; published online 2 March 2016

INTRODUCTION

Hyperbranched and highly branched copolymers have attracted much attention for their unique structures, such as globular shape, void-containing shape and a large number of terminal groups, that lead to unusual properties, such as no crystallization, high solubility and low solution viscosity.^{1–7} Many kinds of hyperbranched and highly branched copolymers have been prepared from AB_x -type monomers not only by step-growth polymerization but also by chain polymerization.

We have also developed initiator-fragment incorporation radical copolymerization (IFIRC) as a novel type of radical copolymerization for the convenient one-pot synthesis of hyperbranched copolymers (Scheme 1).^{8–21} The key point to produce hyperbranched copolymers is the use of a high concentration of radical initiator relative to divinyl monomers. In general, radical copolymerization in the presence of a divinyl monomer results in gelation to yield an insoluble, crosslinked copolymer, the molecular weight of which is considered to be extremely high or infinite.^{22–24} On the contrary, the use of a much higher initiator concentration in the radical copolymerization can cause so great a decrease in the molecular weight that the resulting copolymer with limited molecular weight becomes soluble and has a branched structure.

The thus obtained copolymers contain a large number of initiator fragments as terminal groups that are incorporated via initiation and primary radical termination. Their chemical compositions can be determined from integral intensities of the representative signals of individual monomeric units in the ^1H nuclear magnetic resonance (NMR) spectra. However, the composition determination becomes more difficult as the composition of the branching unit increases because of the signal broadening that occurs as a consequence of the slowing down of the molecular motion, and thereby results in the severe overlap of signals assigned to different monomeric units.

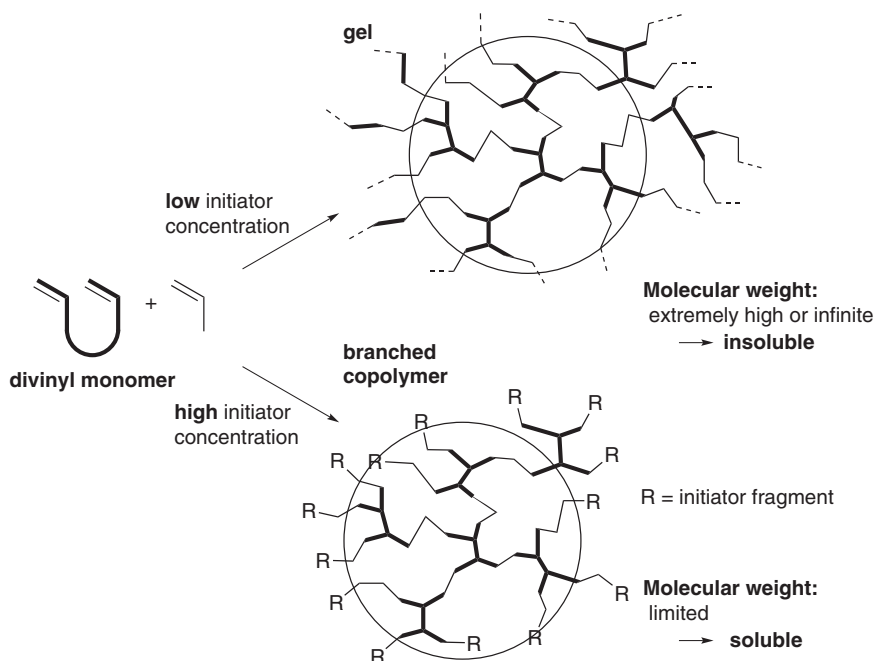
Recently, we have reported that multivariate analysis of NMR spectra is useful for structural analysis of synthetic (co)polymers.^{25–27} For example, principal component analysis (PCA) of ^{13}C NMR spectra of copolymers of methyl methacrylate and *tert*-butyl methacrylate (TBMA) with various chemical compositions, the homopolymers of the two methacrylates and blends of the homopolymers with various blend ratios successfully extracted information of not only chemical composition but also monomer sequence, without assigning the individual signals.^{25,26} Furthermore, chemical compositions of the copolymers were practically predicted by partial least-squares (PLS) regression of ^{13}C NMR spectra of the homopolymers of the two methacrylates, and the blends of the homopolymers.²⁵

¹Department of Chemical Science and Technology, Institute of Technology and Science, Tokushima University, Tokushima, Japan and ²Corporate Research Laboratories, Mitsubishi Rayon Co., Ltd, Hiroshima, Japan

Correspondence: Dr T Hirano, Department of Chemical Science and Technology, Institute of Technology and Science, Tokushima University, 2-1 Minamijosanjima, Tokushima 770-8506, Japan.

E-mail: hirano@tokushima-u.ac.jp

Received 24 November 2015; revised 27 December 2015; accepted 6 January 2016; published online 2 March 2016



Scheme 1 Concept of the initiator-fragment incorporation radical copolymerization (IFIRC) to obtain soluble branched copolymers.

Accordingly, we investigated multivariate analysis of NMR spectra of copolymers having branching structure to examine the extent to which multivariate analysis is applicable to extracting structural information from the NMR spectra, in which signals broaden with increase in the branching-unit composition. The branched copolymers were prepared by IFIRC of ethylene glycol dimethacrylate (EGDMA) and TBMA with dimethyl 2,2'-azobisisobutyrate (MAIB).

EXPERIMENTAL PROCEDURE

Materials

EGDMA (supplied by Mitsubishi Rayon Co., Ltd, Otake, Japan) was purified by washing 1 N aqueous NaOH, followed by drying with MgSO_4 . TBMA (supplied by Mitsubishi Rayon Co., Ltd) was distilled under reduced pressure. MAIB (supplied by Otsuka Chemical Co., Ltd, Tokushima, Japan) was recrystallized from methanol. *N,N*-dimethylformamide (Kanto Chemical Co., Inc., Tokyo, Japan), methanol and lithium bromide (LiBr, anhydride) (Kishida Chemical Co., Ltd, Osaka, Japan) were used without further purification.

Copolymerization

In a typical copolymerization procedure, EGDMA (0.299 g, 1.51 mmol), TBMA (0.427 g, 3.00 mmol) and MAIB (1.043 g, 4.53 mmol) were diluted with 20 ml *N,N*-dimethylformamide in 100 ml round bottom flask. The solution was degassed by several freeze–pump–thaw cycles. Polymerization was carried out at 80 °C under a nitrogen atmosphere. After 3 h, the polymerization mixture was cooled to room temperature and poured into a large volume of methanol/water mixtures (3/7–5/5 vol/vol). The polymer precipitate was collected by centrifugation and dried *in vacuo*. The copolymer yield was determined gravimetrically. The yield was estimated based on the total weight of EGDMA, TBMA and MAIB*, because a large number of initiator fragments were incorporated as terminal groups in the resulting copolymers. Note that the weight obtained by subtracting N_2 from that of MAIB was used as the weight of MAIB*, taking account of N_2 elimination in the MAIB decomposition.

Size-exclusion chromatography measurement

The molecular weights and molecular weight distributions of the polymers were determined by size-exclusion chromatography; the chromatograph was calibrated with standard poly(methyl methacrylate) samples. Size-exclusion

chromatography was performed on an HLC 8220 chromatograph (Tosoh Corporation, Tokyo, Japan) equipped with TSKgel columns (SuperHM-M (6.5 mm inner diameter \times 150 mm) and SuperHM-H (6.5 mm inner diameter \times 150 mm); Tosoh Corporation). *N,N*-dimethylformamide containing 10 mmol l^{-1} LiBr was used as an eluent at 40 °C and a flow rate of 0.35 ml min^{-1} . The initial polymer concentration was set at 1.0 mg ml^{-1} .

Multivariate analysis of ^{13}C NMR spectra

The polymer samples were dissolved in chloroform-*d* (8% wt/vol). The ^1H and ^{13}C NMR spectra of the sample solutions were measured at 55 °C on a ECX400 spectrometer (JEOL Ltd., Tokyo, Japan) equipped with a 10 mm multinuclear probe (^1H : 45° pulse (8.5 ms), pulse repetition 8.90 s, 16 scans; ^{13}C : 45° pulse (7.5 ms), pulse repetition 2.73 s, 5000 scans, with ^1H broadband decoupling). Chemical composition was determined from integral intensities of ^1H NMR signals of the ester groups of the EGDMA, TBMA and MAIB fragment units.

Each ^{13}C NMR spectrum was stored into 32768 complex data points covering a spectral width of 31 250 Hz, and zero-filled to 131 072 points before Fourier transformation. An exponential apodization function was applied to the free induction decays corresponding to a line-broadening factor of 2.0 Hz. The ^{13}C NMR chemical shifts were reference to internal tetramethylsilane (that is, $\delta = 0.0$ p.p.m.).

Bucket integration at an interval of 0.02 p.p.m. was performed with JEOL Alice2 ver.5 for metabolome ver.1.6 software for the resonance regions of the carbonyl carbons; 172–180 p.p.m. Sum of integrated intensities was normalized to 100. Average integrated intensity was subtracted from individual integrated intensity to obtain mean-centered bucket integrated values. PCA of the thus obtained data sets was carried out using Alice2 ver.5 for metabolome ver.1.6 software. PLS regression of the data sets composed of the spectral matrix and primary structure data matrix was conducted using Pattern Recognition Systems Sirius ver.7.0 software (Pattern Recognition Systems, Bergen, Norway). The data were subjected to leave-one-out cross-validations, followed by PLS-2 analysis, in which the chemical compositions of the EGDMA monomeric unit, TBMA monomeric unit and MAIB fragment were simultaneously predicted. The NIPALS (nonlinear iterative partial least squares) algorithm was used for the multivariate analyses.

RESULTS AND DISCUSSION

Preparation of polymer samples by IFIRC

For preparation of branched copolymers with different compositions, IFIRC was carried out in *N,N*-dimethylformamide at 80 °C with a fixed concentration of the vinyl groups. Taking account of bifunctionality of the EGDMA monomer, the concentrations of each monomer were calculated by the following equation;

$$[\text{EGDMA}]_0 \times 2 + [\text{TBMA}]_0 = 0.3 \text{ mol l}^{-1}$$

The concentration of 3.0 mol l⁻¹ TBMA was adopted only in the homopolymerization of TBMA. The radical polymerization of EGDMA in benzene at 80 °C required at least threefold amount of the MAIB initiator relative to the EGDMA monomer to prevent the system from gelation.¹⁴ The concentration of the initiator was therefore calculated by the following equation;

$$[\text{MAIB}]_0 = [\text{EGDMA}]_0 \times 3 + [\text{TBMA}]_0 \times 0.005$$

Generally, a few mol% of initiator relative to monomer is employed in homopolymerization of monovinyl monomers. However, a larger amount of initiator was already required to preclude the gelation. The 0.5 mol% initiator relative to the TBMA monomer was therefore calculated. Pendant vinyl groups disappeared from the poly(EGDMA) prepared by the polymerization for 3 h, whereas a small amount of pendant vinyl groups remained in poly(EGDMA) prepared by the 15 min polymerization. The polymerization time was therefore set to be 3 h. Total 14 polymer samples including 2 homopolymers and 12 copolymers with different compositions were prepared as summarized in Table 1. The two homopolymers were abbreviated as H-0 and H-100, in which the number corresponds to the percent composition of TBMA monomeric unit. Similarly, the 12 copolymers were abbreviated as C-8 to C-94.

PCA to extract information on primary structures

PCA was performed on the data set for the C=O region of 17 samples, composed of 2 homopolymers, 12 copolymers and 3 homopolymer blends. The three blends were prepared by mixing H-0 and H-100 with different blend ratio, as summarized in Table 1. The contribution rates for the first principal component (PC1) and second principal component (PC2) were 87.7% and 10.0%, respectively. The first two principal component factors accounted for 97.7% of the spectral information of the data set. Figure 1 shows the PCA loading plots with corresponding NMR spectra of poly(EGDMA), poly(EGDMA-*co*-TBMA) and poly(TBMA) (cf. H-0, C-52 and H-100 in Table 1). The PCA loadings are the eigenvectors of the cross-product matrix of the spectral space. They therefore contain information on the spectral variations of the original data set.

Sharp positive PC1 loadings were observed at the same chemical shifts of the signals of poly(TBMA), whereas broad negative PC1 ones were observed at the region corresponding to the broad signals of branched poly(EGDMA). This result suggests that PC1 reflects the chemical compositions of TBMA and EGDMA units in the polymer samples. On the other hand, sharp positive PC2 loadings were observed not only between the signals of poly(TBMA) but also at 178.5–179.2 p.p.m. The sharp loadings between the signals of poly(TBMA) are assumed to correspond to the signals assignable to the TBMA units adjacent to EGDMA units and MAIB fragments in the copolymers. The loadings around 178.5–179.2 p.p.m. can be assigned to the signals of the MAIB fragments, because these signals disappear in the ¹³C NMR spectrum of poly(EGDMA-*co*-TBMA) prepared with 2,2'-azobisisobutyronitrile, instead of MAIB (Supplementary Figure S1). In addition, sharp negative PC2 loadings at the same chemical shifts of the signals of poly(TBMA) were observed along with

Table 1 Preparation of polymer samples: homopolymers, copolymers, blends and test samples

Code	Composition in feed ^a /mol%			Composition in copolymer ^b /mol%			Yield/%	M _n ^c × 10 ⁻⁴	M _w /M _n ^c	DB ^d
	EGDMA ^e	TBMA	MAIB	EGDMA	TBMA	MAIB fragment				
H-0	50.0 (100)	0.0	150.0	56.3	0.0	43.7	30	14.2	5.7	1.00
C-8	45.0 (90)	10.0	135.1	49.8	8.1	42.1	33	9.3	4.7	0.92
C-17	40.0 (80)	20.0	120.1	43.2	16.9	39.9	19	9.1	3.5	0.83
C-26	35.0 (70)	30.0	105.2	36.6	25.6	37.8	38	7.1	2.3	0.74
C-35	30.0 (60)	40.0	90.2	31.8	35.1	33.1	31	5.8	1.9	0.65
C-37	25.0 (50)	50.0	75.3	30.7	37.3	32.0	32	5.1	1.7	0.63
C-52	20.0 (40)	60.0	60.3	21.1	52.1	26.8	38	4.5	1.5	0.48
C-63	15.0 (30)	70.0	45.4	17.5	62.6	19.9	51	4.2	1.4	0.37
C-68	10.0 (20)	80.0	30.4	11.5	68.3	20.2	89	3.7	1.3	0.32
C-77	7.5 (15)	85.0	22.9	8.5	77.0	14.5	63	3.9	1.3	0.23
C-83	5.0 (10)	90.0	15.5	5.9	83.3	10.8	61	3.9	1.3	0.17
C-88	2.5 (5)	95.0	8.0	3.7	88.1	8.2	73	4.5	1.4	0.12
C-94	1.5 (3)	97.0	4.4	1.3	94.4	4.3	81	5.5	1.4	0.06
H-100 ^f	0.0 (0)	100.0	0.5	0.0	99.3	0.7	81	19.0	5.7	0.00
B-27 ^g	—	—	—	41.2	26.6	32.2	—	—	—	0.73
B-52 ^g	—	—	—	26.8	52.0	21.2	—	—	—	0.48
B-76 ^g	—	—	—	13.1	76.2	10.7	—	—	—	0.24
test ^h	—	—	—	22.8	53.7	23.5	—	—	—	—

Abbreviations: DB, degree of branching; EGDMA, ethylene glycol dimethacrylate; MAIB, dimethyl 2,2'-azobisisobutyrate; TBMA, *tert*-butyl methacrylate.
^a[EGDMA]₀ × 2 + [TBMA]₀ = 0.3 mol l⁻¹, [MAIB]₀ = [EGDMA]₀ × 3 + [TBMA]₀ × 0.005, in *N,N*-dimethylformamide (DMF) at 80 °C for 3 h.

^bDetermined from ¹H nuclear magnetic resonance (NMR) signals.

^cDetermined by size-exclusion chromatography (SEC; poly(methyl methacrylate) standards).

^dCalculated with DB = (F_{EGDMA} + F_{MAIB}) / (F_{EGDMA} + F_{TBMA} + F_{MAIB}).

^eThe values in the parentheses correspond with the composition of the vinyl groups.

^f[TBMA]₀ = 3.0 mol l⁻¹.

^gBlends of H-0 and H-100.

^hBlends of C-17 and C-88.

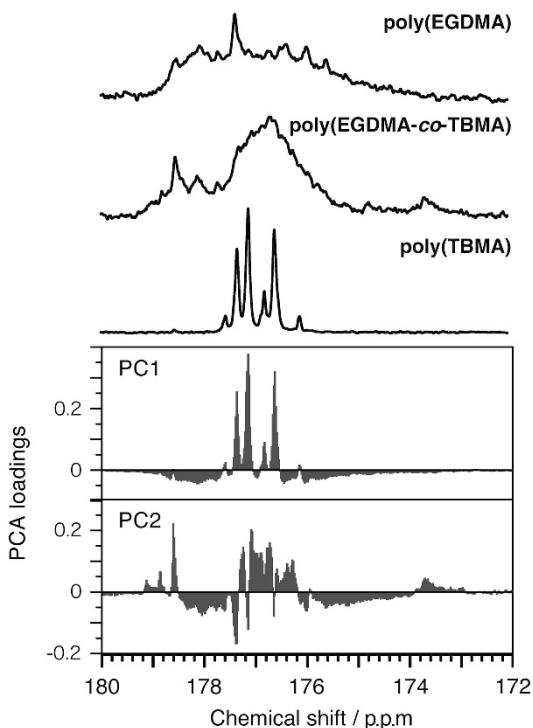


Figure 1 The first principal component (PC1) and second principal component (PC2) loading plots with the corresponding nuclear magnetic resonance (NMR) spectra of the carbonyl carbons of poly(EGDMA), poly(EGDMA-*co*-TBMA) and poly(TBMA) (cf. H-0, C-52 and H-100 in Table 1). EGDMA, ethylene glycol dimethacrylate; PCA, principal component analysis; TBMA, *tert*-butyl methacrylate.

broad negative PC2 ones. These results suggest that PC2 reflects the monomer sequences and initiator fragments in the copolymer samples.

Karhunen–Loève plots for the PC1 and PC2 scores are shown in Figure 2. The poly(EGDMA), poly(TBMA) and their blends (H-0, H-100, B-27, B-52 and B-76) were plotted on a straight line parallel to the PC1 axis. The PC1 scores increased linearly with increase in the TBMA composition. On the other hand, the copolymers prepared by IFIRC were plotted on an asymmetrical parabolic line. These results are consistent with the results that PC1 and PC2 reflect the chemical composition and monomer sequence, respectively, as described above. Strictly speaking, however, the poly(EGDMA) is not homopolymer but copolymer, owing to the MAIB fragment incorporation. Similarly, the poly(EGDMA-*co*-TBMA) is terpolymer. Such a structural feature of the branched copolymers prepared by IFIRC would cause the asymmetric plots for the copolymers; relative mol ratio of the EGDMA unit and MAIB fragment of the branched copolymers varies depending on the copolymerization conditions, whereas that of the blend is unvarying. Nevertheless, the well-ordered plots indicate that PCA successfully extracts information of the primary structures of the branched copolymers even from their NMR spectra that contain not only sharp signals but also broadened signals.

Training set for PLS regression

PLS regression was applied to the data set obtained by bucket integration of the ^{13}C NMR signals of the C=O groups to predict the chemical composition in the polymer samples. For calibration, chemical compositions of each sample, separately determined from

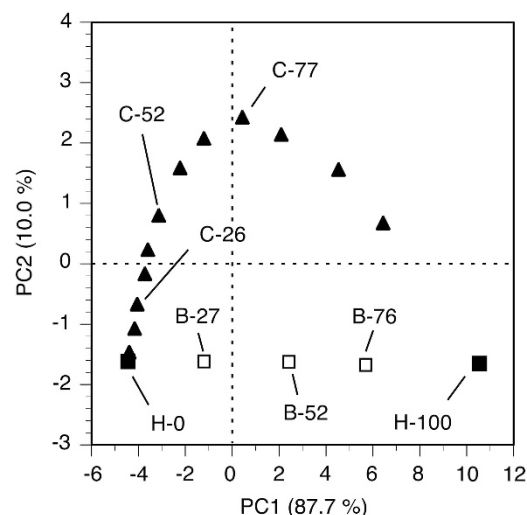


Figure 2 Karhunen–Loève plots for the first principal component (PC1) and second principal component (PC2) scores.

Training set 1:

5 samples (homopolymers + blends)
H-0, H-100, B-27, B-52, B-76

Training set 2:

8 samples (homopolymers + blends + copolymers)
H-0, H-100, B-27, B-52, B-76, C-8, C-52, C-77

Training set 3:

17 samples (homopolymers + blends + copolymers)
H-0, H-100, B-27, B-52, B-76, C-8, C-17, C-26, C-35, C-37, C-52, C-63, C-68, C-77, C-83, C-88, C-94

Scheme 2 Training sets used for predictions by partial least-squares (PLS) regression.

the ^1H NMR spectra, were used. To examine the extent to which the training set affects the prediction, training set 1, composed of poly(EGDMA), poly(TBMA) and three homopolymer blends (H-0, H-100, B-27, B-52 and B-76), was at first chosen as a training set (Scheme 2). This is because the chemical compositions of linear poly(MMA-*co*-TBMA)s can be predicted using a training set composed of the corresponding homopolymers and their blends.^{25,26}

Regression model was constructed with two latent variables (LV1 and LV2), because the predicted compositions of EGDMA unit, TBMA unit and MAIB fragment quite agreed with the corresponding observed composition, determined from ^1H NMR spectra, in leave-one-out cross-validation with two latent variables (LV1 = 99.0% and LV2 = 0.6%) (Supplementary Figure S2). It is noted that, in spite of low contribution rate of the LV2, the use of only LV1 clearly decreases correlation coefficients (R^2) for the leave-one-out cross-validation (Supplementary Figure S2).

PLS regression was therefore conducted to predict the composition of the 12 copolymers (C-8 to C-94) using the training set 1 (Scheme 2) with two latent variables. The predicted values were unexpectedly deviated from the values determined by ^1H NMR

(Figure 3). Regardless of the composition, the compositions of the EGDMA units and MAIB fragments were overpredicted, whereas those of the TBMA units were underpredicted.

Then, the branched copolymers (C-8, C-52 and C-77) were added to the training set 1 (the training set 2 in Scheme 2). Leave-one-out cross-validations with three latent variables (LV1 = 90.9%, LV2 = 7.8% and LV3 = 0.5%) showed good correlation between the compositions determined by ^1H NMR and those predicted by PLS regression

(Supplementary Figure S3). PLS regression with three latent variables was therefore conducted to predict the compositions of other nine copolymers using the training set 2. The predicted values agreed well with the values determined by ^1H NMR, with R^2 of 0.993–0.998 and relative s.d. of 2.9–6.9% (Figure 4).

The PLS loadings clearly indicated the benefit of adding copolymers in the training set (Figure 5). The training sets 1 and 2 gave almost the same LV1 loadings. However, quite different LV2 loadings were

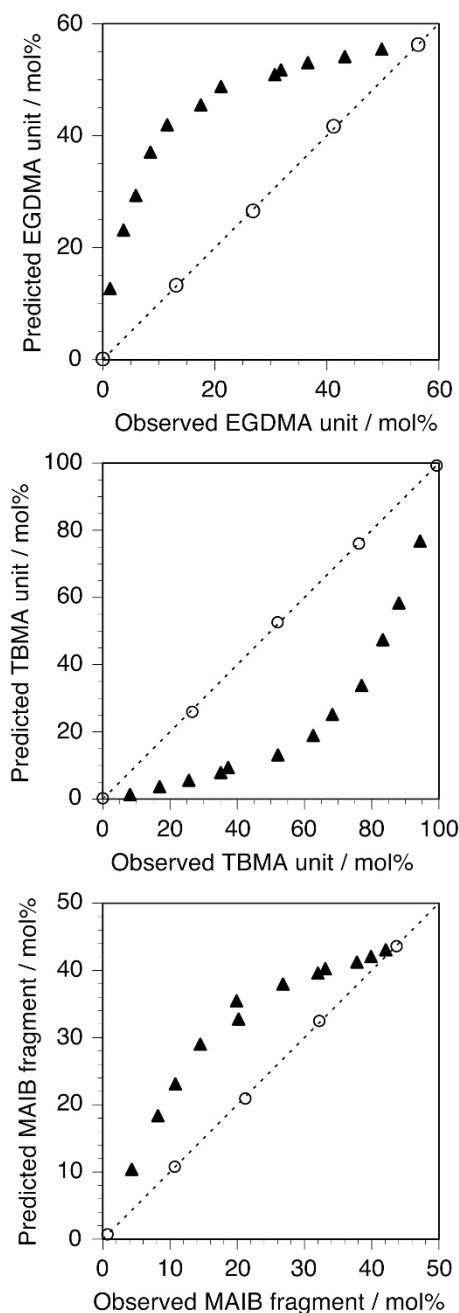


Figure 3 Relationship between the chemical compositions of branched copolymers determined by ^1H nuclear magnetic resonance (NMR) and those predicted by partial least-squares (PLS) regression using the training set 1 with two latent variables. The symbol \blacktriangle denotes plots of target branched copolymers, whereas \circ denotes those of the training set. EGDMA, ethylene glycol dimethacrylate; MAIB, dimethyl 2,2'-azobisisobutyrate; TBMA, *tert*-butyl methacrylate.

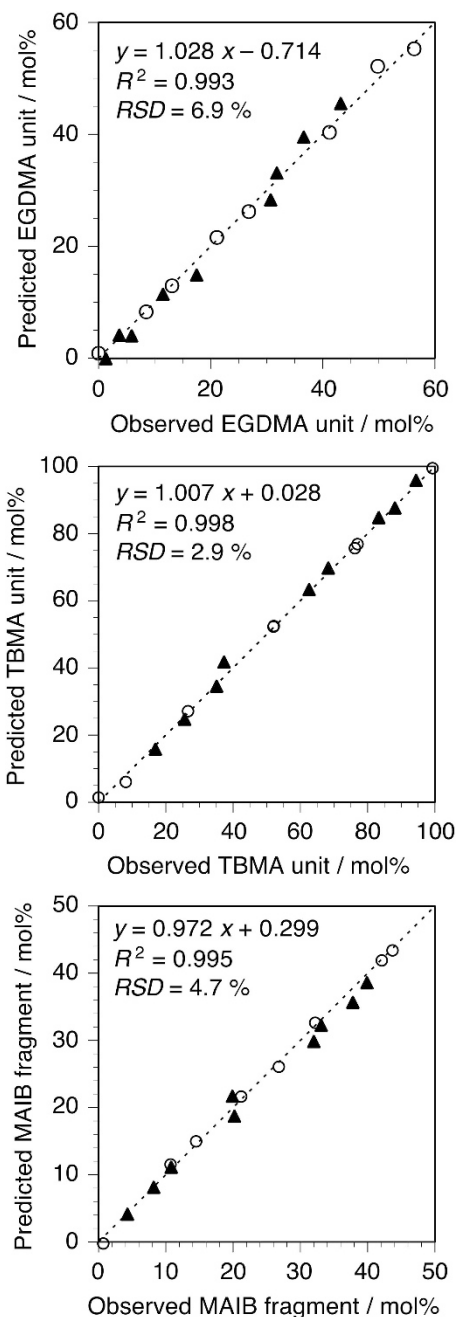


Figure 4 Relationship between the chemical compositions of branched copolymers determined by ^1H nuclear magnetic resonance (NMR) and those predicted by partial least-squares (PLS) regression using the training set 2 with three latent variables. The symbol \blacktriangle denotes plots of target branched copolymers, whereas \circ denotes those of the training set. EGDMA, ethylene glycol dimethacrylate; MAIB, dimethyl 2,2'-azobisisobutyrate; TBMA, *tert*-butyl methacrylate.

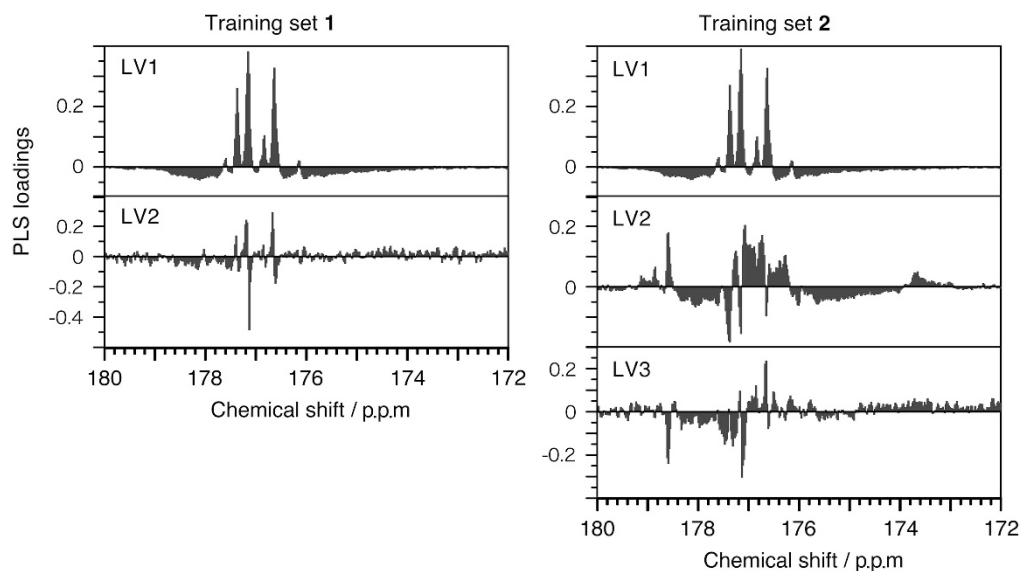


Figure 5 Partial least-squares (PLS) loading plots using the training set 1 with two latent variables and the training set 2 with three latent variables.

observed between the training sets 1 and 2; the information of monomer sequences and initiator fragments were reflected in the LV2 loadings with the training set 2, whereas almost no information on the primary structures was reflected in those with the training set 1. In addition, the training set 2 provided the LV3 loadings having some information on the monomer sequences and initiator fragments. These results suggest that signal broadening of the EGDMA unit likely causes a failure in distinction of monomer sequences, in particular in the signals of the EGDMA unit. It can be therefore concluded that selection of appropriate training set is important for the accurate prediction of the compositions in branched copolymers from the NMR spectra.

Prediction of the DB values by PLS regression

Degree of branching (DB) is one of the important features of branched polymers. By referring to definition of the DB for hyperbranched dendritic polyesters,²⁸ we defined the DB of the branched copolymers obtained by the IFIRC as follows:

$$DB = (F_{EGDMA} + F_{MAIB}) / (F_{EGDMA} + F_{TBMA} + F_{MAIB})$$

where F_{EGDMA} , F_{TBMA} and F_{MAIB} denote the mole fractions of the EGDMA unit, TBMA unit and MAIB fragment, respectively, in the branched copolymers. The DB value of the branched copolymer composed of EGDMA and MAIB fragment is unity, whereas that of the linear polymer of TBMA is close to zero. The DB value is therefore bounded between zero and unity. The DB values were calculated with the compositions in the branched polymers, as summarized in Table 1.

The leave-one-out cross-validation of the DB values using the training set 2 with three latent variables (LV1 = 90.9%, LV2 = 7.8% and LV3 = 0.5%) suggested the accurate prediction of the DB values (Supplementary Figure S4). PLS regression was therefore conducted to predict the DB values of the nine copolymers. The predicted values agreed well with the values calculated from the chemical composition, with an R^2 of 0.998 and the relative s.d. of 3.5% (Figure 6). This means that the DB values can be directly predicted by PLS regression from the NMR spectra even without determining the chemical compositions.

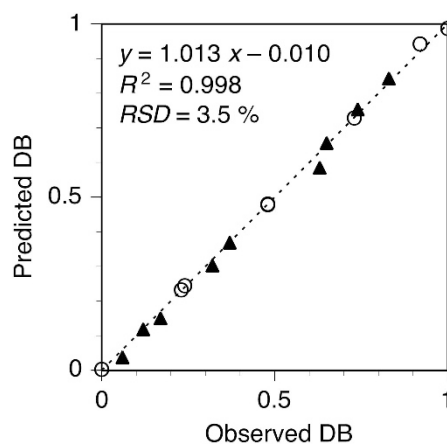


Figure 6 Relationship between the degree of branching (DB) values of branched copolymers calculated from the compositions determined by ^1H nuclear magnetic resonance (NMR) and those predicted by partial least-squares (PLS) regression using the training set 2 with three latent variables. The symbol ▲ denotes plots of target branched copolymers, whereas ○ denotes those of the training set.

Prediction of chemical compositions of unknown sample

To explore the extent to which the PLS regression can be applied to unknown samples, we prepared a test sample (Table 1) by mixing the branched copolymers with different compositions (C-17 and C-88). Compositions of the test sample, calculated based on the weight fractions of the original branched copolymers, are comparable with those of C-52. However, their NMR spectra quite differed from each other (Figure 7).

All the homopolymers, blends and copolymers were selected as a training set (the training set 3 in Scheme 2) for the prediction of the test sample, because a training set containing branched copolymers leads to a more accurate prediction. Indeed, the contribution rates of the second and third latent variables (LV1 = 87.3%, LV2 = 10.4% and LV3 = 0.9%) increased as compared with those with the training set 2.

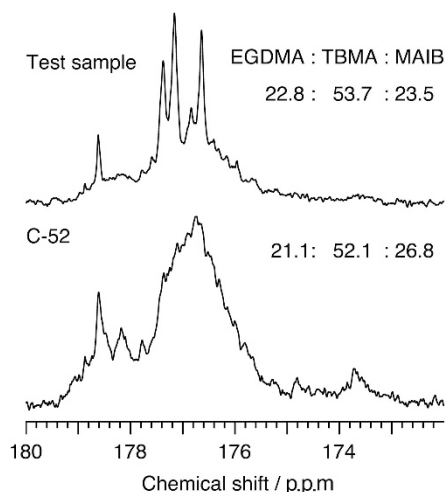


Figure 7 The ^{13}C nuclear magnetic resonance (NMR) spectra of the carbonyl carbons of the test sample and branched copolymer, in the latter of which the chemical compositions are comparable with those in the test sample. EGDMA, ethylene glycol dimethacrylate; MAIB, dimethyl 2,2'-azobisisobutyrate; TBMA, *tert*-butyl methacrylate.

In addition, the high R^2 values were observed in the leave-one-out cross-validations (Supplementary Figure S5). These results suggest successful prediction of the chemical compositions using the training set 3 with three latent variables.

The chemical compositions were predicted to be EGDMA/TBMA/MAIB fragment = 26.1 : 49.8 : 24.1, respectively. Good agreements in the chemical compositions within 3.9% indicates that the multivariate analyses of NMR spectra are practically useful to predict the chemical composition even of the branched polymers, whose NMR spectra are essentially broadened because of the slowing down of the molecular motion.

CONCLUSIONS

Multivariate analyses of ^{13}C NMR spectra of poly(EGDMA), poly(TBMA), their blends and poly(EGDMA-*co*-TBMA)s were found to be useful to practically characterize the primary structures of branched copolymers. PCA successfully extracted information of the primary structures of the branched copolymers from their NMR spectra. It should be noted that NMR spectra of much branched copolymers were drastically broadened as compared with those of less branched copolymers.

The chemical compositions of the branched copolymers were successfully determined by PLS regression without any assignments of ^{13}C NMR signals. It appeared that selection of the training set was important for the prediction of chemical compositions in the branched copolymers: the lack of branched copolymers in the training set resulted in the overprediction of the EGDMA unit. The signal broadening as a consequence of the slowing down of the molecular motion in branched copolymers caused no distinction of monomer sequences in the EGDMA units. In addition, PLS regression using the appropriate training set allowed us to predict not only the chemical compositions but also the DB values.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported in part by KAKENHI (a Grant-in-Aid for Young Scientists (B) (23750130)).

- 1 Kim, Y. H. Hyperbranched polymers 10 years after. *J. Polym. Sci., Part A Polym. Chem.* **36**, 1685–1698 (1998).
- 2 Inoue, K. Functional dendrimers, hyperbranched and star polymers. *Prog. Polym. Sci.* **25**, 453–571 (2000).
- 3 Jikei, M. & Kakimoto, M.-A. Hyperbranched polymers: a promising new class of materials. *Prog. Polym. Sci.* **26**, 1233–1285 (2001).
- 4 Tomalia, D. A. & Fréchet, J. M. J. Discovery of dendrimers and dendritic polymers: a brief historical perspective. *J. Polym. Sci., Part A Polym. Chem.* **40**, 2719–2728 (2002).
- 5 Gao, C. & Yan, D. Hyperbranched polymers: from synthesis to applications. *Prog. Polym. Sci.* **29**, 183–275 (2004).
- 6 Voit, B. Hyperbranched polymers—all problems solved after 15 years of research? *J. Polym. Sci., Part A Polym. Chem.* **43**, 2679–2699 (2005).
- 7 Higashihara, T., Segawa, Y., Sinananwanich, W. & Ueda, M. Synthesis of hyperbranched polymers with controlled degree of branching. *Polym. J.* **44**, 14–29 (2012).
- 8 Sato, T., Sato, N., Seno, M. & Hirano, T. Initiator-fragment incorporation radical polymerization of divinylbenzene in the presence of glyoxylic oxime ether: formation of soluble hyperbranched polymer. *J. Polym. Sci., Part A Polym. Chem.* **41**, 3038–3047 (2003).
- 9 Sato, T., Arima, Y., Seno, M. & Hirano, T. Formation of a soluble hyperbranched polymer via initiator-fragment incorporation radical copolymerization of divinyl adipate and isobutyl vinyl ether. *Polym. Int.* **53**, 1138–1144 (2004).
- 10 Sato, T., Hashimoto, M., Seno, M. & Hirano, T. Soluble hyperbranched polymer through initiator-fragment incorporation radical copolymerization of ethylene glycol dimethacrylate and α -ethyl β -N-(α -methylbenzyl) itaconamate in benzene. *Eur. Polym. J.* **40**, 273–282 (2004).
- 11 Sato, T., Higashida, N., Hirano, T. & Seno, M. Initiator-fragment incorporation radical copolymerization of divinylbenzene and *N*-isopropylacrylamide with dimethyl 2,2'-azobisisobutyrate: formation of soluble hyperbranched polymer nanoparticle. *J. Polym. Sci., Part A Polym. Chem.* **42**, 1609–1617 (2004).
- 12 Sato, T., Ihara, H., Hirano, T. & Seno, M. Formation of soluble hyperbranched polymer through the initiator-fragment incorporation radical copolymerization of ethylene glycol dimethacrylate with *N*-methylmethacrylamide. *Polymer* **45**, 7491–7498 (2004).
- 13 Sato, T., Miyagi, T., Hirano, T. & Seno, M. Initiator-fragment incorporation radical polymerization of ethylene glycol dimethacrylate in the presence of 1,1-diphenylethylene: synthesis and characterization of soluble hyperbranched polymer nanoparticles. *Polym. Int.* **53**, 1503–1511 (2004).
- 14 Hirano, T., Ihara, H., Miyagi, T., Wang, H., Seno, M. & Sato, T. Formation of soluble hyperbranched polymer nanoparticles by initiator-fragment incorporation radical polymerization of ethylene glycol dimethacrylate. *Macromol. Chem. Phys.* **206**, 860–868 (2005).
- 15 Hirano, T., Tanaka, K., Wang, H., Seno, M. & Sato, T. Effect of nitrobenzene on initiator-fragment incorporation radical polymerization of divinylbenzene with dimethyl 2,2'-azobisisobutyrate. *Polymer* **46**, 8964–8972 (2005).
- 16 Sato, T., Arima, Y., Seno, M. & Hirano, T. Initiator-fragment incorporation radical polymerization of divinyl adipate with dimethyl 2,2'-azobisisobutyrate: kinetics and formation of soluble hyperbranched polymer. *Macromolecules* **38**, 1627–1632 (2005).
- 17 Hirano, T., Higashida, N., Wang, H. & Sato, M. S. T. Synthesis and characterization of soluble hyperbranched polymer via initiator-fragment incorporation radical polymerization of divinylbenzene with dimethyl 2,2'-azobisisobutyrate. *J. Appl. Polym. Sci.* **100**, 664–670 (2006).
- 18 Sato, T., Nakamura, T., Seno, M. & Hirano, T. Soluble hyperbranched copolymer via initiator-fragment incorporation radical copolymerization using a trivinyl monomer. *Polymer* **47**, 4630–4637 (2006).
- 19 Sato, T., Nobutane, H., Hirano, T. & Seno, M. Hyperbranched acrylate copolymer via initiator-fragment incorporation radical copolymerization of divinylbenzene and ethyl acrylate: synthesis, characterization, hydrolysis, dye-solubilization, Ag particle-stabilization, and porous film formation. *Macromol. Mater. Eng.* **291**, 162–172 (2006).
- 20 Sato, T., Nomura, K., Hirano, T. & Seno, M. Initiator-fragment incorporation radical polymerization of diallyl phthalate: kinetics, formation of hyperbranched polymer, and iridescent porous film thereof. *J. Appl. Polym. Sci.* **102**, 408–415 (2006).
- 21 Sato, T., Nomura, K., Hirano, T. & Seno, M. Hyperbranched polymer through initiator-fragment incorporation radical copolymerization of divinyl adipate with allyl acetate: synthesis, characterization, dye solubilization, metal-nanoparticle stabilization, and porous film formation. *Polym. J.* **38**, 240–249 (2006).
- 22 Flory, P. J. Molecular size distribution in three dimensional polymers. I. gelation. *J. Am. Chem. Soc.* **63**, 3083–3090 (1941).
- 23 Stockmayer, W. H. Theory of molecular size distribution and gel formation in branched-chain polymers. *J. Chem. Phys.* **11**, 45–55 (1943).
- 24 Stockmayer, W. H. Theory of molecular size distribution and gel formation in branched polymers II. general cross linking. *J. Chem. Phys.* **12**, 125–131 (1944).

- 25 Momose, H., Hattori, K., Hirano, T. & Ute, K. Multivariate analysis of ^{13}C NMR spectra of methacrylate copolymers and homopolymer blends. *Polymer* **50**, 3819–3821 (2009).
- 26 Momose, H., Maeda, T., Hattori, K., Hirano, T. & Ute, K. Statistical determination of chemical composition and monomer sequence distribution of poly(methyl methacrylate-*co-tert*-butyl methacrylate)s by multivariate analysis of ^{13}C NMR spectra. *Polym. J.* **44**, 808–814 (2012).
- 27 Hirano, T., Anmoto, T., Umezawa, N., Momose, H., Katsumoto, Y., Oshimura, M. & Ute, K. Application of multivariate analysis of NMR spectra of poly(*N*-isopropylacrylamide) to assignment of stereostructures and prediction of tacticity distribution. *Polym. J.* **44**, 815–820 (2012).
- 28 Hawker, C. J., Lee, R. & Frechet, J. M. J. One-step synthesis of hyperbranched dendritic polyesters. *J. Am. Chem. Soc.* **113**, 4583–4588 (1991).

Supplementary Information accompanies the paper on Polymer Journal website (<http://www.nature.com/pj>)