

Determination of the Formaldehyde-to-Phenol Molar Ratios of Resol Resins by Infrared Spectroscopy and Multivariate Analysis

Hilkka Holopainen,¹ Leila Alvila,¹ Tuula T. Pakkanen,¹ Jouni Rainio²

¹Department of Chemistry, University of Joensuu, P.O. Box 111, FIN-80101 Joensuu, Finland

²Bakelite Oy, FIN-82430 Puhos, Finland

Received 9 July 2002; accepted 27 December 2002

ABSTRACT: The formaldehyde/phenol (F/P) ratios of resol resins were successfully predicted by the recording of infrared (IR) spectra of both calibration and analyzed resins and by a multivariate analysis technique. In the creation of applicable models, the best correlating IR spectral areas were found between 1800 and 700 cm^{-1} . The positive effects of the increasing replicates and the omission of first-derivative preprocessing on model quality were proven by systematic testing. The characteristic statistical parameters were

acceptable when the resin was similar to the calibration resins. Although the calibration samples had narrow F/P molar ratios (2.00–2.40), or a particular urea content or alkalinity, the best calibration model could also successfully predict the F/P molar ratios of resins with greater F/P ratios, higher urea contents, and lower alkalinity. © 2003 Wiley Periodicals, Inc. *J Appl Polym Sci* 89: 3582–3586, 2003

Key words: resins; FT-IR; infrared spectroscopy

INTRODUCTION

Infrared (IR) spectroscopy is a useful method for analyzing phenol–formaldehyde resol resins. Spectra can be recorded economically and rapidly, and so an IR analysis is worth studying along with statistical methods. The multivariate analysis of near-infrared spectral data has successfully been used, for example, to determine nitrogen, starch, and carbohydrate contents of pine needles¹ and mixed oxygenate concentrations in gasoline.² Multivariate analysis has also been used in Fourier transform infrared (FTIR) spectroscopy of coals,^{3,4} wood lignins,⁵ crude oil resin fractions,⁶ and major components of lignocellulose.⁷ Also, our previous comprehensive study consisted of the IR analysis and chemometric quantitative analysis of phenol–formaldehyde resins.⁸

The aim of this study was to develop a multivariate method of determining formaldehyde/phenol (F/P) molar ratios of phenol–formaldehyde resins rapidly and accurately by IR spectroscopy during resin condensation or immediately after condensation. The method could be used to control the F/P molar ratio of a resin product as part of quality control.

EXPERIMENTAL

Resin synthesis

To create a model and to test its prediction ability, 13 phenol–formaldehyde resol resins (Table I) were pre-

pared by Dynoresin Oy (Bakelite Oy, Puhos, Finland) and stored frozen at -18°C until the IR analysis. Resins 1–7, with F/P molar ratios between 2.00 and 2.40, were used as a calibration set. These F/P ratios were calculated from the amounts of the starting reagents formaldehyde and phenol. At the end of the condensation reaction, 2.2 wt % extra urea was added to resins 1–7 to control the viscosity. Various resins (8–13) with special characteristics were used for prediction.

IR equipment

IR spectra of the resins were measured with a PerkinElmer System 2000 FTIR spectrometer equipped with a Golden Gate Single Reflection Diamond ATR P/N 10500 series sample device. This accessory could be used to record IR spectra for solids, liquids, and films. The Golden Gate accessory used a 2 mm \times 2 mm diamond, which was brazed into a tungsten carbide disc.⁹ A sample, one drop of a well-mixed liquid resin without any extra solvent or matrix, was placed on the surface of the diamond.

The resolution of the IR spectra of both the calibration and prediction resins was 2.00 cm^{-1} . The number of scans was 16; apodization was set to strong, and the optical path difference velocity was set to 0.5 cm/s . All spectra were recorded at room temperature within 30 s so that the effect of evaporation would be avoided.

After the IR recording of each resin, the surface of the diamond was carefully cleaned with methanol; this ensured the repeatability and reliability of the results with-

Correspondence to: L. Alvila.

TABLE I
Condensation Conditions and Properties of the Resins for Calibration and Prediction

Resin	Special characteristic	F/P molar ratio in Synthesis	Viscosity (mPas)	Catalyst	Alkalinity (wt %)	Urea (wt %)	Solid content (wt %)	Free formalin (wt %)
For calibration								
1		2.00	476	NaOH	6.40	2.2	46.8	0.01
2		2.10	530	NaOH	6.80	2.2	47.3	0.01
3		2.15	535	NaOH	6.20	2.2	47.0	0.02
4		2.20	462	NaOH	6.40	2.2	48.1	0.05
5		2.25	530	NaOH	6.40	2.2	49.1	0.11
6		2.30	486	NaOH	6.40	2.2	48.1	0.18
7		2.40	620	NaOH	6.45	2.2	49.8	0.47
For prediction								
8	Similar to the calibration resin 2	2.10	500	NaOH	6.45	2.2	48.9	0.01
9	High urea content	2.10	308	NaOH	6.15	8.3	50.0	0.02
10	Low alkalinity	2.10	432	NaOH	5.50	2.2	47.4	0.05
11	Low solid content	2.10	428	NaOH	5.62	2.2	42.6	0.02
12	High F/P	3.33	32	NaOH		0.0	52.9	4.84
13	High F/P and the TEA catalyst	3.40	23	TEA		4.0	48.6	2.73

out outliers. Such spectra should be high-quality and as identical as possible from the same sample.

IR software

The software consisted of IR Data Manager Release 2 and a chemometric program called Quant+ from PerkinElmer. The IR data manager was used to acquire, modify, and print spectra. The Quant+ program, a chemometric quantitative analysis software package specially suited to a multicomponent spectroscopic analysis, changed the IR spectrum into data points and calculated with statistical methods the correlations between the IR spectrum and the original data set (i.e., the F/P molar ratios calculated from the amounts of the starting reagents formaldehyde and phenol). The calibration spectra were treated as a matrix and subjected to principal component analysis. Correlations were established by multiple linear regression. The Quant+ program created a model that could be used for prediction.^{8,10} IR spectra were processed in the absorbance mode in the Quant+ program.

RESULTS AND DISCUSSION

Models

The first three models, 1–3, were built to determine the proper number of replicates needed to achieve the optimum model (Table II). Each of the seven standard resins, 1–7, were recorded with IR once (model 1), twice (model 2), or three times (model 3). The standard error of estimate (SEE) gave an indication of the quality of fit of the regression, and the standard error of prediction (SEP) was the magnitude of the error expected when independent samples were predicted with the model.¹⁰ These standard errors decreased with increasing replicates. An *F* value, being a measure of the signal to noise, and a variance, the proportion of variability, should be as high as possible for a good model.¹⁰ In this model optimization, all these values were the highest in the case of three replicates. Therefore, an increasing number of replicates was found to improve the precision of the model. Because small differences were found between the spectra of consecutive runs of even the same resin sample, and because the IR analysis was fast and quite economic,

TABLE II
Effect of the Number of Replicates on Models for the Determination of F/P Molar Ratios

Model	Spectral area (cm ⁻¹)	Number of replicates	SEE	SEP	<i>F</i> value	Variance (%)	Number of PCs
1	4000–600	1	0.044	0.053	50**	90.84	2
2	4000–600	2	0.037	0.044	141***	92.13	1
3	4000–600	3	0.018	0.027	247***	98.41	4

PC, principal component.
 ** *p* < 1%.
 *** *p* < 0.1%.

TABLE III
Effect of Spectral Area on Models for the Determination of F/P Molar Ratios

Model	Spectral area (cm ⁻¹)	SEE	SEP	F value	Variance (%)	Number of PCs
Effect of the OH area						
3	4000–600	0.018	0.027	247***	98.41	4
4	2900–600	0.012	0.012	516***	99.23	4
Effect of other than the OH area						
5	2400–600	0.009	0.014	951***	99.58	4
6	2200–600	0.009	0.014	871***	99.54	4
7	1880–600	0.010	0.014	845***	99.53	4
8	1700–600	0.008	0.014	1100***	99.64	4
9	1580–600	0.012	0.013	577***	99.31	4
10	1540–600	0.012	0.013	746***	99.47	4
11	1500–600	0.009	0.013	1025***	99.61	4
12	1400–600	0.024	0.026	265***	96.72	2
13	1300–600	0.027	0.026	214***	95.97	2
14	1800–600	0.009	0.014	942***	99.58	4
15	1800–700	0.005	0.007	1798***	99.87	6
16	1800–800	0.010	0.014	753***	99.47	4
17	1800–900	0.008	0.007	1391***	99.71	4
Effect of the area with strongest correlation						
18	1800–1295, 1283–700	0.006	0.008	1962***	99.85	5
19	1529–1195, 1052–850	0.006	0.008	1557***	99.85	6
20	1529–1195, 1052–985	0.005	0.006	2335***	99.90	6
21	1529–1400, 1052–985	0.008	0.009	1552***	99.64	3
22	1529–1400	0.006	0.009	1414***	99.84	6
23	1457–1195, 1052–985	0.002	0.003	6401***	99.98	9
24	1580–1195	0.006	0.010	1455***	99.87	7
25	1580–1460, 1360–1195, 1052–985	0.004	0.005	3230***	99.93	6

*** $p < 0.1\%$.

the best accuracy was worth achieving with at least three replicates.

The IR signals could be assigned with the previous results⁸ and literature.^{11–14} To obtain the most suitable spectral area, we tested models 3–25 (Table III). The favorable effect of the omitted OH IR area of 4000–2900 cm⁻¹ was found at lower SEE and SEP values and at higher F and variance values. The unfavorable effect of the area (4000–2900 cm⁻¹) may be due to the signals of water [$\nu(\text{OH}) \sim 3500$ cm⁻¹], formic acid [$\nu(\text{OH}) = 3300$ – 2500 cm⁻¹], and urea [$\nu(\text{NH}) = 3500$ – 3300 cm⁻¹], which overlapped the $\nu(\text{OH})$ signals of the methylol and phenolic groups of the resins.

Furthermore, the effect of the spectral regions within 2400–600 cm⁻¹ was tested by the selection of the particular data ranges in the interval of 40–320

cm⁻¹ (models 5–17, Table III). The best results—the minimum SEE value of 0.005, the minimum SEP value of 0.007, the highest F value of 1798, and the highest variance of 99.87%—were reached when the spectral area was 1800–700 cm⁻¹. The data ranges correlating best with the property of interest (the F/P molar ratio) could be observed as positive bands in the property correlation plot. The strongest correlations were found at 1500 [$\nu(\text{C}=\text{C})$ of phenol], 1475 [aliphatic d(CH₂) of methylol], 1445 [$\nu(\text{C}=\text{C})$ of phenol], 1250 [$\nu_{\text{ip}}(\text{C}-\text{O})$ of phenol], and 1015 cm⁻¹ [$\nu(\text{C}-\text{O})$ of methylol]. Therefore, these specific areas were selected as the basis of data ranges of models 18–25 (Table III). The best statistical results (low errors of 0.002–0.006, high F values of 2335–6401, and high variances of 99.90–99.98%) were found with models 20, 23, and 25.

TABLE IV
Effect of First Derivative on Models for the Determination of F/P Molar Ratios

Model	Spectral area (cm ⁻¹)	First derivative smoothing points	SEE	SEP	F value	Variance (%)	Number of PCs
15	1800–700	— ^a	0.005	0.007	1798***	99.87	6
26	1800–700	5	0.016	0.025	618***	98.56	2
27	1800–600	19	0.018	0.022	996***	98.13	1
28	1800–600	49	0.008	0.010	1525***	99.63	3

^a First derivative preprocessing was not used.

*** $p < 0.1\%$.

TABLE V
New Calibration of the Models for the Determination of F/P Molar Ratios

Model	Spectral area (cm ⁻¹)	SEE	SEP	F value	Variance (%)	Number PCs
15'	1800–700	0.006	0.007	1621***	99.83	5
20'	1529–1195, 1052–985	0.003	0.005	4266***	99.96	7
23'	1457–1195, 1052–985	0.008	0.009	896***	99.69	5
28' ^a	1800–600	0.009	0.011	1435***	99.63	3

^a First derivative preprocessing was used.
*** $p < 0.1\%$.

A first derivative is often used for preprocessing data to remove the undesired structure and to reduce the effects of noise.¹⁵ In this study, the effect of using the first derivative on model validity was tested for models 26–28 as well as model 15, in which the first derivative was not used at all (Table IV). According to Wold and Sjöström,¹⁵ preprocessing with the first derivative could remove some data relevant to the modeling and prediction. Also, in our study, using the derivative seemed to have a negative influence. However, more valid models were also obtained with first-derivative preprocessing when a smoothing degree of a curve was increased with more smoothing points.

Four models—15', 20', 23', and 28' (Table V)—were recreated by the recording of new calibration spectra and recalibration. Model 20' was the best one, with

minimum errors of 0.003–0.005, a very valid *F* value of 4266, and an extremely high variance of 99.96%.

Prediction of the F/P molar ratios of the resins

We used the best models to study their ability to predict first the F/P molar ratio of resin 8, which was similar to calibration resin 2 and second the F/P values of resins 9–13 with at least one special characteristic. The F/P molar ratios of these six phenol–formaldehyde resol resins (8–13) were predicted by optimized models 15', 20', 23', and 28' (Table VI).

The root-mean-square (RMS) error gives an overall indication of the quality of reproduction of an observed spectrum. An RMS error of a good estimate is lower than 0.081;¹⁰ on this basis, all four tested models

TABLE VI
Prediction of F/P Molar Ratios^a

Resin	Model	Statistical parameters				F/P molar ratio			
		RMS error	Peak to peak	<i>M</i> distance	<i>F</i> value	In synthesis	Predicted	<i>R</i> error	<i>M</i> distance
8	15'	0.002	0.018	0.9	10**	2.10	2.05	0.01	0.8
	20'	0.001	0.008	15.6	72***		2.08	0.01	5.7
	23'	0.002	0.013	4.5	119***		2.00	0.01	2.2
	28'	0.001	0.007	0.5	15**		2.04	0.01	0.7
9	15'	0.012	0.061	57.6	700***	2.10	2.02	0.03	48.1
	20'	0.003	0.015	25.5	391***		2.08	0.01	23.5
	23'	0.003	0.014	19.4	199***		2.10	0.02	10.9
	28'	0.002	0.011	2.5	106***		1.97	0.01	1.8
10	15'	0.002	0.017	4.6	17**	2.10	2.07	0.01	3.9
	20'	0.001	0.005	17.9	46***		2.09	0.01	6.7
	23'	0.002	0.012	7.7	84***		2.02	0.02	7.7
	28'	0.001	0.006	1.3	20***		2.07	0.01	0.6
11	15'	0.004	0.033	3.0	82***	2.10	2.00	0.01	2.6
	20'	0.002	0.012	53.9	227***		2.03	0.01	5.7
	23'	0.004	0.028	8.0	419***		1.98	0.01	3.4
	28'	0.001	0.011	0.7	43***		2.02	0.01	0.6
12	15'	0.023	0.131	285.7	2467***	3.33	3.16	0.07	238.0
	20'	0.014	0.066	400.7	8880***		3.30	0.05	270.0
	23'	0.012	0.063	351.7	3477***		2.88	0.12	333.3
	28'	0.006	0.036	29.9	1327***		2.47	0.02	17.0
13	15'	0.013	0.105	178.3	811***	3.40	2.81	0.06	148.7
	20'	0.008	0.044	204.3	2707***		2.87	0.03	130.3
	23'	0.009	0.043	157.7	1870***		2.72	0.08	163.3
	28'	0.004	0.026	27.3	675***		2.52	0.02	12.6

^a The averages of the prediction results for three measurements of every resin.
** $p < 1\%$.
*** $p < 0.1\%$.

were very valid for every resin. The peak-to-peak error, giving an indication of localized deviations, should be no more than about five times the RMS error.¹⁰ For resin 9, this condition was realized with all the models.

When a sample is highly compatible with the calibration set, the Mahalanobis (*M*) distance ratio is below 1.0.¹⁰ The lowest total *M* distances and the lowest *M* distances for the calculated F/P molar ratios were naturally found in the case of resin 2, which was similar to resin 8 included in the calibration set. The other *M* distance values were higher than the normal values, revealing that the modified resins (high urea contents or F/P ratios, low alkalinity or solid contents, and different catalysts) were different from the ones in the range of the calibration set.¹⁰ The *F* value (the *F* ratio or residual ratio being defined as the ratio of the variance of the errors between the original and calculated spectra) should be even below 3.0 when features of unknown samples are particularly successfully modeled in calibration.¹⁰ The lowest *F* values were found in the prediction of resin 8. The *F* value depended on the selected spectral area and increased when narrow data ranges were used. Although the *M* distance and the *F* value should be small, higher values seemed to have no significant effect on the final predicted F/P molar ratios of the studied resins.

The best model, 20', gave a predicted F/P molar ratio of 2.09 for resin 10 with low alkalinity with one replicate and naturally gave an equally good ratio with three replicates. The prediction could be regarded as successful (F/P in synthesis = 2.10) when the confidence interval, the *R* error, of 0.01 was taken into account. Also, for resin 12 with a high F/P ratio, the *R* error of 0.05 obtained with model 20' was high enough to increase the predicted F/P molar ratio of 3.30 to the level of the F/P molar ratio in synthesis (3.33). This is worth noticing because the F/P ratios were totally different from those of the calibration resins, and usually in a multivariate analysis, an evenly distributed calibration set is a basic requirement for a good model. Furthermore, a high urea content did not disturb the prediction of the F/P molar ratio. Model 20' gave the same valid predictions of 2.08 ± 0.01 for resins 8 and 9, both having an F/P value of 2.10 in synthesis. When resin 11 had a low solid content or when the catalyst was triethylamine (TEA; resin 13) instead of NaOH, which was used in

the calibration resins, the results were not acceptable even with the best model, 20'.

Model 23' was accurate in the prediction of the F/P value of resin 9, which had a much higher urea content (8.3 wt %) than that (2.2 wt %) of the calibration resins. When preprocessing with the first derivative was included in the prediction model 28', the results deviated more from the F/P values used for resin synthesis than with the other models.

CONCLUSIONS

The F/P molar ratios of phenol-formaldehyde resol resins can be determined with IR spectroscopy and multivariate analysis. The spectra for model construction and for prediction have to be recorded under identical conditions and with several replicates. The best correlating IR spectral areas and the effect of preprocessing (e.g., the use of the first derivative) can be determined with comprehensive testing. The best models are successful in the quantitative prediction of F/P molar ratios, even after the resins to be predicted are moderately modified. Naturally, a wider and more representative calibration set covering all major variations would be worth studying.

References

1. Hiukka, R. *Chemom Intell Lab Syst* 1998, 44, 395.
2. Faber, N. M.; Duewer, D. L.; Choquette, S. J.; Green, T. L.; Chesler, S. N. *Anal Chem* 1998, 70, 2972.
3. Fredericks, P. M.; Lee, J. B.; Osborn, P. R.; Swinkels, D. A. J. *Appl Spectrosc* 1985, 39, 303.
4. Fredericks, P. M.; Lee, J. B.; Osborn, P. R.; Swinkels, D. A. J. *Appl Spectrosc* 1985, 39, 311.
5. Faix, O.; Böttcher, J. H. *Holzforschung* 1993, 47, 45.
6. Midttun, O.; Sjöblom, J.; Kvalheim, O. M. *Prog Colloid Polym Sci* 1998, 108, 175.
7. Schultz, T. P.; Templeton, M. C.; McGinnis, G. D. *Anal Chem* 1985, 57, 2867.
8. Holopainen, T.; Alvila, L.; Rainio, J.; Pakkanen, T. T. *J Appl Polym Sci* 1998, 69, 2175.
9. Golden Gate Single Reflection Diamond ATR P/N 10500 Series User Manual; Specac: Smyrna, GA, 1998.
10. Quant+ User's Manual; PerkinElmer: Beaconsfield, England, 1991.
11. Bagghi, D.; Malakar, R. K. *Paintindia* 1985, 25, 19.
12. Erä, V. A.; Salo, H.; Kaps, T.; Lindberg, J. J. *Angew Makromol Chem* 1975, 48, 185.
13. Erä, V. A.; Lindberg, J. J.; Mattila, A.; Vauhkonen, L.; Linna-halme, T. *Angew Makromol Chem* 1976, 50, 43.
14. Ebewele, R. O.; River, B. H.; Koutsky, J. A. *J Appl Polym Sci* 1986, 31, 2275.
15. Wold, S.; Sjöström, M. *Chemom Intell Lab Syst* 1998, 44, 3.