ELSEVIER

Journal of Chromatography A, 756 (1996) 89–106

JOURNAL OF CHROMATOGRAPHY A

Ruggedness testing of a size-exclusion chromatographic assay for low-molecular-mass polymers

Y. Vander Heyden[a], C. Hartmann[a], D.L. Massart[a,*], A.M.J. Hollands[b], P. Nuyten[b], P. Schoenmakers[b]

[a]*ChemoAC, Pharmaceutical Institute, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussel, Belgium*
[b]*Shell International Chemicals B.V. Amsterdam, Badhuisweg 3, 1031 CM Amsterdam, The Netherlands*

Received 14 May 1996; revised 23 July 1996; accepted 23 July 1996

Abstract

A ruggedness test of a size-exclusion chromatographic assay for a low-molecular-mass polymer mixture was performed by applying a fractional factorial design. Eight factors selected from the method procedure were examined at three levels by reflecting the design. The effects of the factors were calculated and interpreted, statistically as well as graphically. The statistical interpretation method based on the use of two-factor interaction effects to estimate experimental error was found to be effective to indicate significant effects in a ruggedness test. The factors column manufacturer and detector type turned out to be the least robust. Variations in the other factors, within the levels examined, do not lead to chromatographically relevant changes.

Keywords: Ruggedness; Fractional factorial design; Polymers

1. Introduction

When developing a size-exclusion-chromatographic (SEC) method, or indeed any analytical method that is to be used at multiple locations, a balance needs to be found between standardisation and flexibility. A method is more likely to yield comparable results at different locations if the column dimensions, the packing material and the manufacturer are tightly specified. However, there are practical advantages associated with the possibility of applying different columns within a given, flexible method. At the other extreme, a completely standardised method that requires a specific brand and type of pump, injector, column oven, detector, etc., is certainly unattractive.

Often problems in a method become apparent during interlaboratory performance studies. At this stage there is an understandably strong desire to make the method pass the tests and an equally strong inclination not to make any significant changes to the method. The only alternative is then to try and improve the reproducibility by extensive or even excessive standardisation.

The great benefits of systematic ruggedness testing are that the chance of a method failing an (intra- or interlaboratory) reproducibility test can be drastically reduced, while excessive standardisation can be avoided in cases where it is not necessary.

Ruggedness testing is a part of method validation. In a ruggedness test the influence of small variations

* Corresponding author.

0021-9673/96/$15.00
PII S0021-9673(96)00633-4

in factors selected from the method description or from environmental parameters on the responses of the method are evaluated [1–4]. The variations (levels) of the factors are selected in a range which represents the variation which can be expected when the method is performed by different analysts and/or on different instruments and/or in different laboratories, etc.

In this study a ruggedness test was performed by executing a one sixteenth-fraction factorial design for eight factors, 2^{8-4}(IV) (see Appendix A). The design was reflected [1–3] to examine the factors at three levels. Because it was of resolution IV [5] the two-factor interactions are not confounded with main effects. The significance of effects was determined (i) statistically using the two-factor interaction effects [2,6,7] to estimate experimental error and (ii) visually by drawing normal probability plots.

The SEC assay studied concerns a low-molecular-mass polymer mixture. The effects of the factors were determined on a number of quantitative responses used to characterise the molecular-mass distribution of the polymer mixture.

This study was also performed to evaluate whether the strategy where two-factor interaction effects are used to estimate the experimental error in a fractional factorial design as applied in Ref. [2] for an HPLC method is effective for an SEC assay as well.

After statistical interpretation, it was checked if a standardisation for the factors found to be statistically significant could be sufficient to obtain chromatographically consistent results.

2. Theory

Synthetic polymers do not consist of molecules with only one molecular mass. They are so called hetero disperse and therefore one determines molecular-mass distributions (MMD). This is done by determining certain average molecular masses as will be explained further. A number of these quantitative responses were measured during the ruggedness test.

For each set of conditions a polystyrene calibration curve is established. In order to do so, two standard solutions containing several monodisperse polystyrenes (PS) with known molecular masses are injected. The top molecular mass (M_{top}) is the value corresponding to the maximum of a polymer peak in the chromatogram. From the chromatograms of the standard solutions, the retention times corresponding to the M_{top} of the peaks can be established, resulting in a calibration curve. The MMD of the sample can be calculated from the chromatogram of the sample solution using the calibration curve. In our case the sample contains several oligomer peaks. These are characterised by calculating the molecular mass corresponding to the peak maxima denoted as MM1, MM2, and MM3. The molecular masses of sample components are calculated based on the assumption that the relationship between size in solution and molecular mass is the same as for PS. When PS calibration curves are used to obtain the MMD of chemically different polymers, as in the present case, we speak of molecular masses relative to PS, as opposed to absolute molecular masses. An example of a chromatogram of the sample is shown in Fig. 1.

The following average molecular masses are described, (a) the number average molecular mass, M_n; (b) the weight average molecular mass, M_w and (c) the z-average molecular mass, M_z. These masses are represented by the following equations:

$$M_n = \frac{\sum N_i M_i}{\sum N_i} \tag{1}$$

$$M_w = \frac{\sum m_i M_i}{\sum m_i} = \frac{\sum N_i M_i^2}{\sum N_i M_i} \tag{2}$$

$$M_z = \frac{\sum m_i M_i^2}{\sum m_i M_i} = \frac{\sum N_i M_i^3}{\sum N_i M_i^2} \tag{3}$$

where N_i is the total number of molecules with molecular mass M_i and m_i is the total mass of all molecules with molecular mass M_i. The average molecular masses M_n, M_w and M_z are measures for respectively the brittleness of the polymer, its tensile strength and its rigidity. The ratios M_w/M_n and M_z/M_w are also determined. M_w/M_n is a measure of the width of the MMD and M_z/M_w of the elastic behaviour of the polymer. The factors selected from the analytical procedure were examined at three levels (see Table 1), the nominal (specified) and two extremes, in a reflected one sixteenth-fraction factorial design. These factors are variables specified in

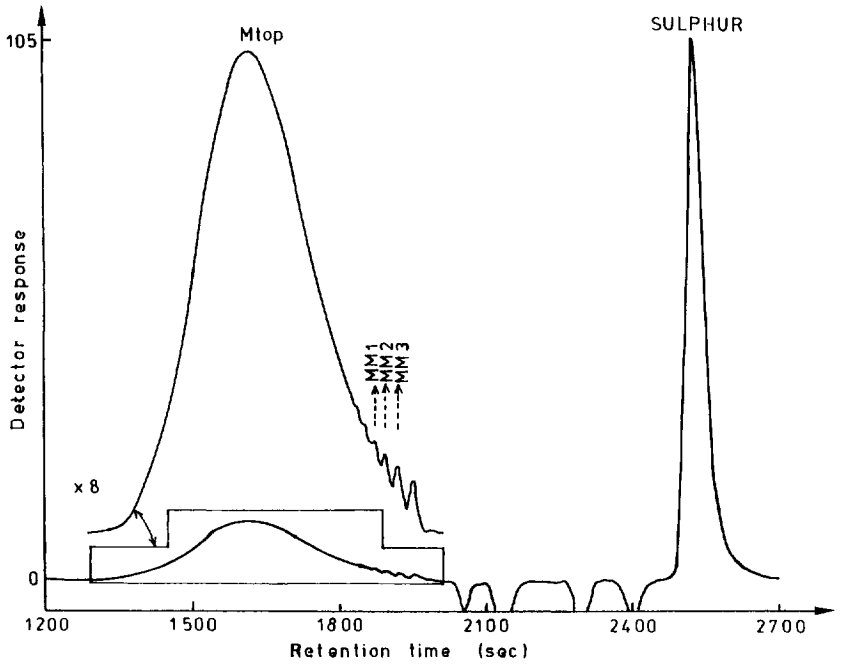

Fig. 1. Elution curve of the sample, showing the main polymer (M_{top}), the oligomers (MM1, MM2, MM3) and sulphur as internal standard. Conditions are nominal.

the procedure but which can vary slightly when the method is applied in another lab than the initial one, e.g., in another lab of the same company. Some theoretical aspects of the applied design are described in Appendix A.

The effects, E_X, of the factors were calculated according to the formulas generally applied [2,3,8,9] and normalised relative to the nominal response, Y_n [2,3].

$$E_X = \frac{\sum Y(\text{highest level})}{N/2} - \frac{\sum Y(\text{lowest level})}{N/2} \quad (4)$$

$$\%E_X = \frac{E_X}{Y_n} \cdot 100 \quad (5)$$

Statistically significant effects were determined using a t-test [2,3,10,11]. The standard error on E_X, $(SE)_e$, is estimated from two-factor interaction effects ($E_{X_iX_j}$) and represents the experimental error in the design:

$$t = \frac{|E_X|}{(SE)_e} \quad (6)$$

with

$$(SE)_e = \sqrt{\frac{\sum E_{X_iX_j}^2}{n_{X_iX_j}}} \quad (7)$$

where $n_{X_iX_j}$ is the number of two-factor interaction effects used to estimate $(SE)_e$. Effects are statistically significant if the t-value calculated for an effect is larger than a tabulated critical t-value, $t_{n_{X_iX_j}}$ with $n_{X_iX_j}$ degrees of freedom or if the $|E_X|$-value is larger than a critical effect ($E_{critical}$ or $\%E_{critical}$) value:

$$|E_X| \geq E_{critical} = t_{n_{X_iX_j}} \cdot (SE)_e \quad (8)$$

or

$$|\%E_X| \geq \%E_{critical} = \frac{E_{critical}}{Y_n} \cdot 100 \quad (9)$$

The two-factor interaction effects were used in the statistical interpretation to estimate the experimental error in the design because in previous studies [2,7] it was observed that these interaction effects were

Table 1
The factors and their levels as examined in the ruggedness test

Factor	Level	Values
A: Flow mobile phase	−1	0.7 ml/min
	0	0.8 ml/min
	+1	0.9 ml/min
B: Column temperature	−1	27°C
	0	22°C
	+1	32°C
C: Injection volume	−1	50 μl
	0	60 μl
	+1	70 μl
D: Injection concentration	−1	2.0 mg/ml
	0	2.5 mg/ml
	+1	3.0 mg/ml
E: Column manufacturer	−1	Phenomenex
	0	PolymerLabs
	+1	Waters Styragel
F: Injection temperature	−1	27°C
	0	22°C
	+1	32°C
G: Detector cell temperature	−1	35°C
	0	30°C
	+1	40°C
H: Detector type (only two levels examined)	−1	Waters 410
	0	ERMA 7512
	+1	Waters 410

Level (0): nominal level; levels (−1) and (+1): extreme levels.

not significant in ruggedness tests on HPLC methods.

Effects were also evaluated by drawing normal probability plots for the normalised effects. These normal probability plots allow us to identify effects which are not normally distributed around zero (significant effects) but also to check the hypothesis that the two-factor interactions are not significant and therefore can be used in the above given statistical interpretation.

3. Experimental

3.1. Nominal chromatographic conditions

The chromatography is performed on a set of three columns with a length of 30 cm and an internal diameter of approximately 7.5 mm, each containing 5 μm particles with a pore size of 10^4, 10^3 and 10^2 Å respectively. The mobile phase is tetrahydrofurane (THF) which is pumped over the columns at a flow-rate of 0.8 ml/min. The column temperature is 22°C and the polymer components are detected with a refractive index detector thermostated at a temperature of 30°C.

3.2. Test solutions at nominal conditions

The sample solution is a 2.5 mg/ml solution of a low molecular-mass polymer mixture in THF. The standard solutions are polystyrene mixtures in THF. Two standard solutions were applied. The first mixture contained polystyrenes with molecular masses of 380 000, 96 000, 22 000, 5050 and 1320, while the second one had masses of 156 000, 49 900, 11 600, 2950 and 580. The signal of the last standard allows individual oligomer peaks down to a molecular mass of 266 to be used for calibration purposes. Elemental sulphur (1 mg per ml THF) is used as an internal standard to correct for flow-rate variations between injections of sample and standards. The concentration of each polystyrene in the mixtures was 1.5 mg/ml. The injection volume is 60 μl and the nominal temperature of the injector is 22°C.

3.3. Materials

The sample polymer mixture and the standard mixtures were available in the laboratory (Shell, Amsterdam, The Netherlands). The SEC equipment consisted of a WISP717 auto-injector (Waters), a Waters 590 pump (Waters), a CROCO-CIL column oven (Cluzeau Info-Lab, France) and an ERC3310 degasser (ERMA). Columns of three different manufacturers were used: (a) Phenomenex (300×7.8 mm I.D.) Phenogel 5 μm particle size, pore sizes 10^4, 10^3, 10^2 Å; (b) PolymerLabs (300×7.5 mm I.D.) PL-GEL 5 μm particle size, pore sizes 10^4, 10^3, 10^2 Å and (c) Waters (300×7.8 mm I.D.) Styragel 5 μm particle size, pore sizes HR4, HR3 and HR1. The refractive index detectors used were an ERMA 7512 and a Waters 410 detector. Between the injector and the columns an in-line filter (SSI) of 0.5 μm was placed.

4. Results and discussion

4.1. Ruggedness experiments

The ruggedness test is meant to evaluate the performance of the method under conditions that are slightly different from the nominal values. A ruggedness test can, for practical reasons, seldom be exhaustive. The number and types of parameters (factors) that can be studied are subjected to restrictions. Eight factors were selected to be examined in the fractional factorial design. These factors and their levels are shown in Table 1.

It is important to study variations (levels) that are realistic and controllable. It is realistic to assume that the actual flow-rate delivered by HPLC pumps varies from instrument to instrument. For example, if the set (expected) flow-rate is 0.8 l/min, the actual one may be 0.85 ml/min. In the case of flow-rate, we estimate that such systematic variations are confined to the range 0.7 to 0.9 ml/min for a nominal flow-rate of 0.8 ml/min. In other words, we assume a maximum variation of about 10%. Likewise, we have considered reasonable variations in other parameters in our experimental design. While such variations may seem trivial, it is of great practical value to know, for example, whether a method performs well within a concentration range of 2–3 mg/ml, or whether a solution of 2.5±0.01 mg/ml needs to be carefully prepared.

We decided to study the column temperature and the injection temperature independently, because they may well be different in practice and because it has been suggested that this may have a significant effect in liquid chromatography [12]. For the factors column temperature, injection temperature and detector cell temperature only deviations above the nominal level were considered to be interesting (e.g., also to avoid solubility problems which can occur at lower temperatures). Therefore the low levels for these factors are of a higher temperature value than the nominal ones.

Some other parameters that could have been included were only identified after the completion of the present test. One example is the effect of extra-column band broadening, which we may have studied indirectly by including different detectors. Extra-column band broadening may be tested systematically by using tubing of different internal diameters.

On the other hand, parameters that cannot be adequately controlled by the person performing a ruggedness test cannot be included as factors in the design. One obvious example is the quality of the eluent. For example, the quality of the THF used in SEC experiments may affect the stability of the baseline and, consequently, the obtained values for the average molecular masses. Also the presence of contaminants in the eluent may affect the adsorption behaviour of the column and thereby the calibration of a SEC system. Unfortunately, it is very difficult to know a priori whether the content of a bottle is good or bad, as this is affected by a large number of factors (manufacturer, batch, age, time after opening, etc.) and the causal relationships are unknown. Therefore it is equally difficult to simulate the effect of the THF quality as part of a ruggedness test.

Short-term variations in the flow-rate, viz. variations within a run, between individual runs or between the analysis and the calibration, are obviously detrimental to the quality of the results obtained from SEC experiments. It is important that such effects are carefully considered during the validation and application of a method. At the validation stage, short-term variations are reflected in the repeatability of a method. Sensible testing of the precision of a method must always proceed in the order (i) repeatability test, (ii) ruggedness test, and (iii) reproducibility test. It is a big waste of time and effort to test the ruggedness of a method that is not sufficiently repeatable. This is even more true if a non-robust method is subjected to an (interlaboratory) reproducibility test. During the application of a method, control charts are very useful to monitor the performance. The elution time of a flow marker, such as elemental sulphur (see Fig. 1), is a meaningful parameter in this respect.

Because our goal was to investigate the ruggedness of a method rather than the effects of short-term fluctuations in e.g., the flow-rate, we recorded calibration curves at each different set of conditions. In other words, we assume that the method can be made to work with good repeatability at a given location.

The design selected to examine the eight factors was the sixteenth-fraction factorial design given in Table 7. To examine the factors at three levels the

design was reflected [1,2] (see Appendix A). For the calculation of effects and the statistical interpretation, the results are treated as coming from two separate designs.

Three injections at nominal conditions were performed for each set of columns before and after the design experiments. They allow the normalisation of the effects and a check that no drastic changes have occurred in the results during the performance of the design. The design experiments were blocked per column set for practical reasons. When changing experimental conditions between experiments the chromatographic device was conditioned for a period of at least 2 h in order to attain the desired circumstances.

4.2. Discussion of results

Fig. 2 shows an overlay of a large number of calibration curves obtained during the test. Three clusters of curves are clearly discernible. Each of these is associated with one of the three types of column sets used in the study. Clearly, and quite expectedly, the choice of column is the most im-

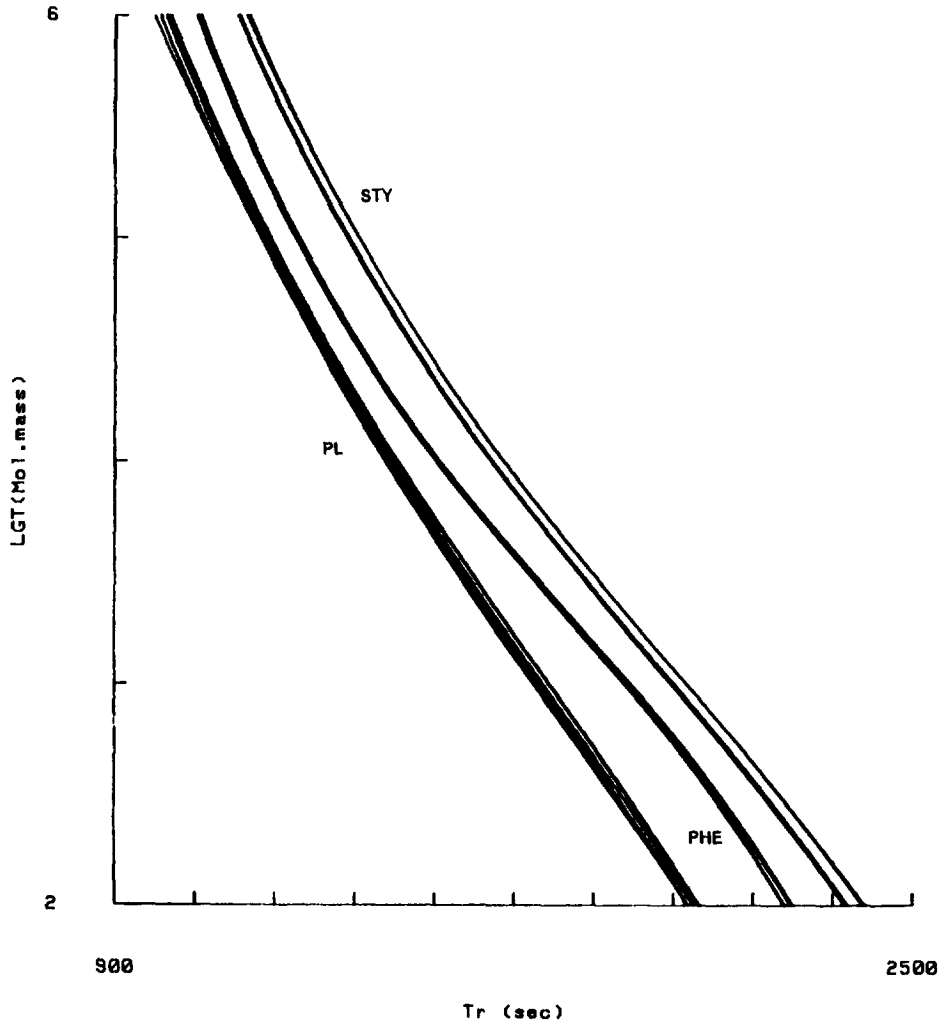

Fig. 2. Plots of the molecular mass as a function of the retention time. The calibration curves are for narrow polystyrene standards in THF. Except for the flow-rate, set at 0.8 ml/min, all other conditions are varied. Three clusters are discernible, each corresponding to a different column set. PHE = Phenomenex; PL = PolymerLabs; STY = Waters Styragel.

portant factor affecting the calibration curve in SEC. Within each cluster, there are smaller, but significant variations due to different values for factors other than the column, especially at the extremes (high and low MM region) of the curves. This may be due to the large influence that small variations in the retention times of the highest and lowest molecular-mass reference standards may have on the curve shape in these parts of the curve. The central part is affected by the retention times of all standards and is thus more robust.

Different SEC columns will give rise to different calibration curves, but in theory they will still yield identical MMDs for any given sample. In practice, there are several reasons why different mass data may result from different calibrations:

(i) Accidental fluctuations in the calibration curves may have an effect. However, this will become apparent during repeatability testing, provided that at this stage the repeatability of the calibration is included.

(ii) Any difference in behaviour between the calibration standards and the sample will affect the outcome. Very often the standards are of different nature (e.g., polystyrene) than the polymer under study. A notorious effect is the adsorption of polymers on the packing material. This will be different for different polymers and it may also be affected by the presence of other materials (solvents, additives) in the sample.

(iii) It is assumed in routine SEC that the chromatographic dispersion (viz. the peak width observed for a material with a specific, invariable molecular mass) is negligible in comparison with the dispersion due to the molecular-mass distribution. This is clearly an approximation and if the band-broadening contribution is significantly different between, for example, different columns, this will affect the molecular-mass data obtained.

The effects and the normalised effects of the factors were calculated for the responses M_{top}, molecular masses of the different oligomers (MM1, MM2, MM3), M_n, M_w, M_z, M_w/M_n and M_z/M_w. The effects were normalised relative to the average response measured at the PolymerLabs column set. The statistical test described above was performed and the $\%E_{critical}$ values were calculated at significance levels of 5% and 1%. In addition to the quantitative responses, the ruggedness of a qualitative response, the specific resolution, was investigated.

All experiments were performed by one analyst but the resulting chromatograms were interpreted independently by two persons. This means that for each effect of a factor two estimations were made. Combining the results of both interpretations leads to a new design with nine factors where the ninth factor is the interpretation by the analyst (see Appendix A). The resulting effects for the first eight factors (A–H) are the average of the effects found from the two separate interpretations.

As illustration the effects and normalised effects of the factors and of the interactions on the molecular mass (M_{top}) of the main peak of the polymer mixture as found by one analyst are shown in Table 2(a) for both parts of the reflected design, as well as the statistical interpretations for this response. In Table 2(b) the results of the combined design are given. The normalised effects obtained by the two analysts separately and those from the combined design were compared in graphs similar to Fig. 3.

For quantitative factors, the effects calculated for the intervals $[-1,0]$ and $[0,+1]$ (Table 2) allow a rough idea about the curvature of the response as a function of the factor levels to be obtained. This can be visualized in so-called effect plots [13] which ease the interpretation of curvature (plots not shown here).

The normal probability plots for the normalised main and interaction effects from Table 2(b) are shown in Fig. 4 for both parts of the design. Effects which are not significantly different from zero tend to fall on a straight line while significant effects deviate from this line. In Fig. 4 this is the case for the effect of factor *E*. The full straight line shown in the figures is the least squares line through all the points. When none of the factors has a clearly significant effect the least squares line fits all points, otherwise the line is attracted to outlying (significant) effects.

The interpretation of the normal probability plots is not based on a specific criterion but on visual inspection. Therefore effects at the limit of significance are sometimes somewhat subjective to interpret. However in this study the normal probability plots are in the first instance not meant to indicate

Table 2
Effects and normalised effects (%Effect) of the factors and of the interactions on the molecular mass (M_{top}) of the main peak of the polymer mixture. The statistical interpretation for that response is also shown

Factor	Interval [−1,0]		Interval [0,+1]	
	Effect	%Effect	Effect	%Effect
(a) *Results from one of the analysts*				
(A) Flow	48.625	1.60	−34.00	−1.12
(B) Column temp.	−101.375	−3.33	70.000	2.30*
(C) Injection volume	1.125	0.04	−3.000	−0.10
(D) Injection conc.	12.625	0.41	46.000	1.51
(E) Column manufacturer	135.625	4.45*	−187.500	−6.15**
(F) Injection temp.	74.125	2.43	55.500	1.82
(G) Detector cell temp.	−7.375	−0.24	−41.500	−1.36
(H) Detector type	−68.375	−2.24	4.500	0.15
Interactions				
AB+CE+DG+FH	23.375	0.77	3.000	0.10
AC+BE+DH+FG	0.875	0.03	17.500	0.57
AD+BG+CH+EF	49.875	1.64	0.500	0.02
AE+BC+DF+GH	80.875	2.65	−26.500	−0.87
AF+BH+CG+DE	35.375	1.16	27.500	0.90
AG+BD+CF+EH	−80.625	−2.64	−6.500	−0.21
AH+BF+CD+EG	−26.625	−0.87	−60.000	−1.97
Statistical interpretation t_{7df} 5%=2.365, 1%=3.499				
	Interval [−1,0]		Interval [0,+1]	
$(SE)_e$	50.7601		27.8164	
Significance level	$E_{critical}$	$\%E_{critical}$	$E_{critical}$	$\%E_{critical}$
5%	120.048	3.94	65.786	2.16
1%	177.610	5.83	97.329	3.19
b) *Results of the combined design (Table 9)*				
(A) Flow	18.563	0.61	−37.688	−1.24
(B) Column temp.	−69.563	−2.28	46.313	1.52
(C) Injection volume	5.688	0.19	−11.063	−0.36
(D) Injection conc.	7.813	0.26	26.438	0.87
(E) Column manufacturer	123.813	4.06**	−175.563	−5.76**
(F) Injection temp.	37.063	1.22	20.313	0.67
(G) Detector cell temp.	4.188	0.14	−18.063	−0.59
(H) Detector type	−43.563	−1.43	12.313	0.40
(T) Analyst-interpretation	−2.406	−0.08	−6.531	−0.21
Interactions				
AB+CE+DG+FH	34.938	1.15	−3.938	−0.13
AC+BE+DH+FG	25.688	0.84	26.188	0.86
AD+BG+CH+EF	19.813	0.65	−2.313	−0.08
AE+BC+DF+GH	43.813	1.44	2.188	0.07
AF+BH+CG+DE	23.563	0.77	9.063	0.30
AG+BD+CF+EH	−48.813	−1.60	33.688	1.11
AH+BF+CD+EG	−22.063	−0.72	−35.438	−1.16
Statistical interpretation t_{7df} 5%=2.365, 1%=3.499				
	Interval [−1,0]		Interval [0,+1]	
$(SE)_e$	32.987		21.328	
Significance level	$E_{critical}$	$\%E_{critical}$	$E_{critical}$	$\%E_{critical}$
5%	78.015	2.56	50.441	1.65
1%	115.423	3.79	74.627	2.45

Intervals [−1,0] and [0,+1] indicate which levels are examined in that part of the reflected design. *=significant at 5% level **=significant at 1% level.

significant and non-significant effects but only as a kind of reference [14] to confirm significance's found with the statistical interpretation and to check if both interpretations do not systematically give different results.

The results of the *t*-test depend on the criterion used to estimate $(SE)_e$ in Eq. (6). The $(SE)_e$ can be estimated from different sources such as replicate measurements at nominal levels, duplicated design experiments, multiple-factor interaction effects,

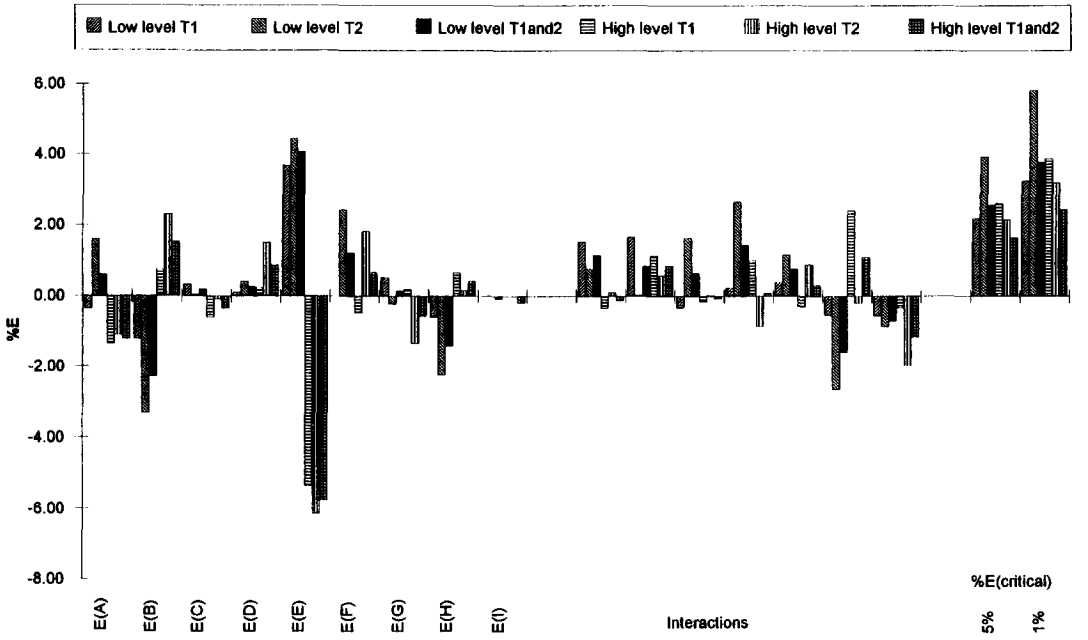

Fig. 3. Comparison of the effects on M_{top} determined by both analysts (T1, T2) and from the combined design (T1 and T2). Interaction and critical effects are also shown.

dummy factor effects or the variance of the design results themselves. It was shown that the number of effects considered statistically significant can be very different depending on the error estimate used [2]. Based on the comparison between the statistical results and the normal probability plots it was found that the use of two-factor interaction effects was appropriate to estimate the experimental error in a fractional factorial design. This was found to be the case for designs performed for ruggedness tests in different application fields as HPLC [2,3,7] and galenics [15].

Sometimes from the normal probability plots it is also possible to obtain information on the existence of anomalous data such as outliers in the measured results [16]. However, the normal probability plots are not very sensitive to indicate possible outliers in design experiments [17]. Nor in our data, was clear evidence found for the occurrence of outliers.

The normal probability plots made for the combined design results show that for all responses all two-factor interaction effects were found in the straight line area of the normal probability plot indicating that they do not significantly differ from zero and can be used in the estimation of the experimental error in the design. The normal probability plots made for the individual results of both analysts showed for some responses a two-factor interaction effect which is situated in the subjective area of limited significance. Based on the results of the combined design this significance was considered negligible and due to the summation of four non-significant two-factor interactions. Each calculated interaction effect is indeed a confounding of four two-factor interaction effects (see Appendix A and Table 2). However when a calculated two-factor effect was situated in the subjective area of limited significance it was mostly the case for the estimation for AG+BD+CF+EH. If the significance of this calculated effect is not due to chance it is probably due to the interaction EH, since this is the only interaction for which the main effect of one of the contributing factors is clearly significant (cf. further, Table 3).

The fact that clearly significant two-factor interactions are absent is not that surprising as at first sight could be thought. Of course, several two-factor interactions are known to be significant in different

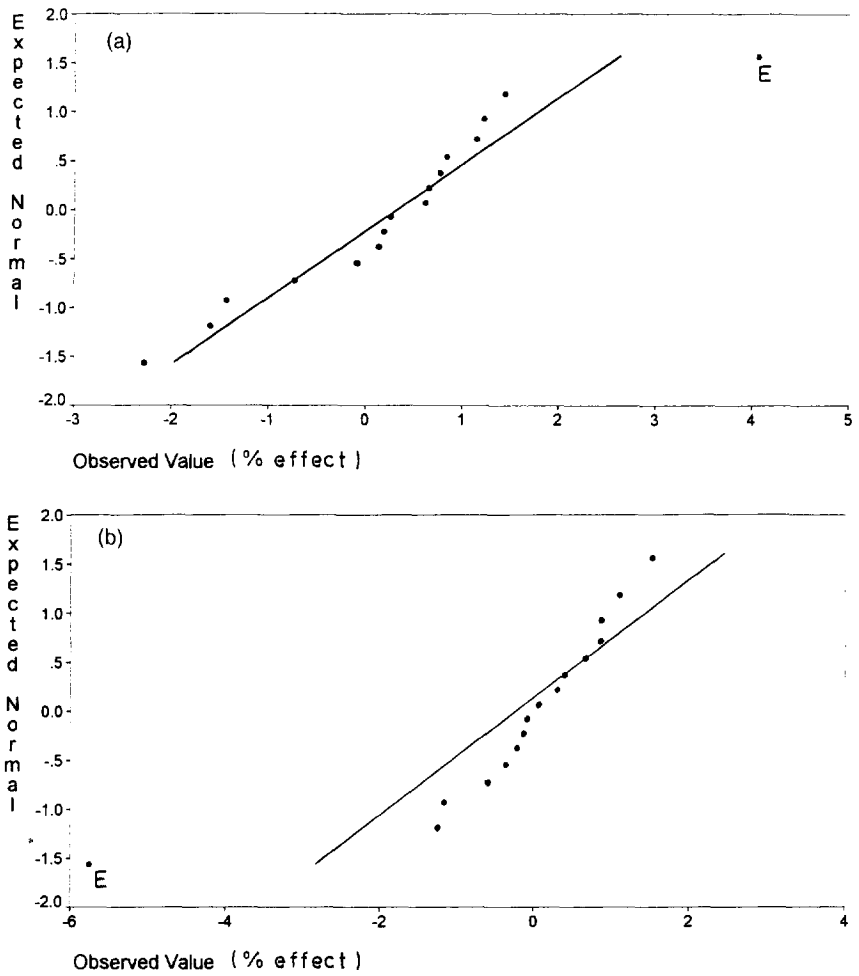

Fig. 4. Normal probability plots for the effects of Table 2(b); (a) from the part of the design with levels (−1) and (0); (b) from the part with levels (0) and (+1).

application fields. However, the study of these interactions is mainly interesting for optimisation purposes when one is building models and response surfaces in order to predict responses and when the factors are examined in much broader intervals, at several levels and by means of other designs than the screening designs applied. We are not dealing with the optimisation step here, since a ruggedness test is part of the validation of a method supposed to be optimised earlier. In a ruggedness test one is basically only interested in indicating which factors can affect the results of the method. In these latter tests the main effects are relatively small since the factors are examined in a narrow interval and most of them are expected to be more or less robust. The two-factor interactions, which represent half of the difference between the effects of a factor X_i at both levels of a factor X_j, will therefore even be considerably smaller and as was seen from these and other results [2,3,7] they are normally not different from the experimental error in a design.

If two-factor interaction effects estimated from screening designs would be significant they indeed cannot be used any more to estimate the experimental error. However, due to the degree of confounding that occurs in those designs such estimated effect is in most cases a confounding of several two-factor interactions (aliases) as was seen higher and it often

Table 3
Summary table of the statistically significant effects from the combined design

Response	Significant factor	Levels	Statistical significance	$\%E_X$	$\frac{\lvert\text{Main effect}\rvert}{\text{critical effect}}$	
					$\alpha=0.05$	$\alpha=0.01$
M_{top}	Column	PL – PHE	>99%	4.06%	1.59	1.07
		STY – PL	>99%	−5.76%	3.49	2.35
MM1	Column	PL – PHE	>99%	−6.31%	5.13	3.47
MM2	Column	PL – PHE	>99%	−5.54%	2.61	1.76
MM3	Column	PL – PHE	>99%	−4.53%	1.94	1.31
M_n	Detector	ERMA – Waters	>99%	−4.75%	1.54	1.04
M_w	Column	PL – PHE	>99%	4.13%	1.96	1.32
		STY – PL	>95%	−1.79%	1.19	(0.81)
M_z	Column	PL – PHE	>95%	6.32%	1.28	(0.86)
	Column temp.	32–22°C	>95%	2.38%	1.05	(0.71)
M_w/M_n	Column	PL – PHE	>99%	6.45%	2.05	1.38
	Detector	ERMA – Waters	>95%	3.31%	1.05	(0.71)
M_z/M_w	Column	STY – PL	>95%	1.52%	1.04	(0.70)
	Detector	Waters – ERMA	>95%	−1.59%	1.09	(0.73)

Column = column manufacturer; PHE = Phenomenex; PL = PolymerLabs; STY = Waters Styragel; ( ) = not significant at α level = 0.01.

will be difficult to draw any conclusion about the different aliases.

Visual and statistical interpretations were again compared as in Refs. [2,3,14,18] in order to examine whether the statistical method applied was also appropriate for this case study in the field of size-exclusion chromatography. Both interpretations, in general, lead again to identical conclusions.

The statistically significant effects of the factors on the different responses are summarised in Table 3. It was observed that the effect of the analyst was significant for none of the responses. Therefore in Table 3 the effects of the combined design which are the average effects found by both analysts are given. Table 3, beside the significant effects $\%E_X$, also shows the statistical significance and the ratio for $|\%E_X|/\%E_{critical}$ for $\alpha=0.05$ and 0.01 to indicate how strongly significant the effects are.

It can be observed that the column manufacturer factor is the least robust (Table 3). PolymerLabs and Phenomenex columns in particular give different results for almost all responses, with the exception of M_n and M_z/M_w. The columns from PolymerLabs and Waters Styragel give more similar results. Only for few responses (M_{top}, M_w and M_z/M_w) a significant effect (different result) is found. These observations were confirmed by the measurements at nominal level with the different columns (see Table 4). Both from the nominal and the design results (effects) it is seen that for the oligomers (MM1, MM2 and MM3) the Phenomenex column gives higher molecular masses than the PolymerLabs column while for the other responses the PolymerLabs column gives higher results. The standard deviations found at nominal levels could be compared between columns and between analysts. Between columns mostly comparable standard deviations were found and, if they were different, the PolymerLabs column gave the lowest. Between analysts it was seen that the standard deviations of T2 for the oligomers at the Phenomenex column are clearly higher than of T1. Also in the design results interpreted by T2 higher $E_{critical}$ values were seen for the oligomers. For the other responses it was observed that higher standard deviations were found by T1 with the Waters Styragel column.

Although, as already mentioned above, different calibration curves obtained with the different columns (Fig. 2) theoretically should lead to the same MMD, it is observed from the above that in practice this is not the case.

Another factor causing a significant effect on some responses is the detector type. While this factor does not influence at all the molecular mass of the peaks,

Table 4
Average responses ($\bar{x}$) and standard deviations (SD) from the nominal level results interpreted by both analysts

Response	Phenomenex column		PolymerLabs column		Waters column	
	T1 ($n=4$)	T2 ($n=6$)	T1 ($n=6$)	T2 ($n=6$)	T1 ($n=6$)	T2 ($n=6$)
Average responses $\bar{x}$						
M_{top}	2875	2798	3049	3049	2846	2849
MM1	497	493	471	471	472	475
MM2	417	410	398	398	401	401
MM3	342	337	327	327	329	329
M_n	1847	1748	1752	1755	1766	1740
M_w	3553	3449	3761	3736	3671	3634
M_z	6289	6132	6986	6829	6799	6654
M_w/M_n	1.923	1.975	2.146	2.128	2.080	2.089
M_z/M_w	1.770	1.778	1.858	1.828	1.852	1.831
Standard deviations (SD)						
M_{top}	52.1	40.5	48.9	48.9	37.2	36.5
MM1	2.2	9.7	3.3	2.3	6.5	4.9
MM2	1.0	6.1	1.7	1.6	4.6	5.1
MM3	3.3	6.0	0.8	0.9	3.0	3.6
M_n	21.3	71.3	21.4	15.2	67.4	25.9
M_w	49.4	50.9	35.4	45.3	46.3	26.5
M_z	156.5	89.6	128.6	185.0	275.6	131.7
M_w/M_n	0.011	0.056	0.014	0.014	0.072	0.030
M_z/M_w	0.022	0.033	0.022	0.029	0.060	0.032

With T1 on the Phenomenex column two outliers could be demonstrated, therefore $n=4$.

it has a significant effect on the responses M_n, M_w/M_n and M_z/M_w. Especially M_n is influenced by this factor. Effects due to the detector could originate from differences in flow cell geometry, band broadening, cell volume and noise levels.

A third factor giving a statistically significant effect was the column temperature. It was found significant on M_z in the interval 22–32°C. This factor was also found to be just below the limit of significance for the responses M_{top}, M_w and M_z/M_w. The other factors that were examined do not seem to have a significant effect on any of the considered responses, at least not in the intervals examined.

The statistical interpretation allows to determine which factors have a significant effect on a response, but for a chromatographer the questions remain whether or not the variation caused by a statistically significant factor is chromatographically relevant. On interpreting the results of the ruggedness test it is assumed that chromatographically acceptable results will be obtained when statistically significant factors are adequately standardized, while the values of the other factors remain and vary within the intervals examined. Therefore it was checked whether or not the variance caused by the different factors in the design, the variance of the design results (s_Y^2), is significantly higher than the variance observed at nominal levels (s_n^2), H_0: $s_Y^2 = s_n^2$; H_1: $s_Y^2 > s_n^2$ and thereupon if this still is the case when one standardises for certain factors (see Table 5). The value used for s_n^2 was the pooled variance observed at nominal levels for the different columns and by the different analysts. It has to be remarked that if for a given method well defined system suitability criteria are used within a laboratory then the s_n^2 could be replaced by a value representing these criteria. For some responses (MM1, M_z, M_z/M_w) the variance from the design results is not significantly higher than s_n^2 although some of them (M_z, M_z/M_w) showed statistically significant effects.

For the other responses H_1 was accepted, i.e., $s_Y^2 > s_n^2$. Thereupon the variance s_Y^2 was considered after standardisation for the most significant factor. For instance when this factor was the column type the variances within the columns were determined from the design results and pooled to give s_Y^2. This

Table 5
Comparison of design (without or with standardisation for one or more factors) variances (s_Y^2) with the variance at nominal levels (s_n^2)

Response	Variance was considered within	$F = \frac{s_Y^2}{s_n^2}$	Conclusion	
			$\alpha = 0.05$	$\alpha = 0.01$
M_{top}				
Low levels ($\%E_{critical} = 2.56\%$)				
	Design	5.36	>[a]	>
	Columns**	3.41	>	>
	Columns + column temperature	2.84	>	>
	Columns + column temperature + detector	2.46	>	=[b]
High levels ($\%E_{critical} = 1.65\%$)				
	Design	6.16	>	>
	Columns**	2.10	>	=
	Columns + column temperature	1.83	=	=
MM1				
Low levels ($\%E_{critical} = 1.23\%$)				
	Design	8.23	>	>
	Columns**	1.20	=	=
High levels ($\%E_{critical} = 1.74\%$)				
	Design	1.79	=	=
MM2				
Low levels ($\%E_{critical} = 2.12\%$)				
	Design	10.83	>	>
	Columns**	3.08	>	>
High levels ($\%E_{critical} = 1.95\%$)				
	Design	3.69	>	>
	Columns	3.30	>	>
MM3				
Low levels ($\%E_{critical} = 2.34\%$)				
	Design	9.02	>	>
	Columns**	4.39	>	>
High levels ($\%E_{critical} = 2.84\%$)				
	Design	6.95	>	>
	Columns	5.65	>	>
M_n				
Low levels ($\%E_{critical} = 3.08\%$)				
	Design	2.18	>	=
	Detectors**	1.34	=	=
High levels ($\%E_{critical} = 4.74\%$)				
	Design	2.78	>	>
	Detectors	2.21	>	=
M_w				
Low levels ($\%E_{critical} = 2.11\%$)				
	Design	6.81	>	>
	Columns**	3.53	>	>
	Columns + injection temperature	3.14	>	>
	Columns + injection temperature + detector	3.21	>	>
High levels ($\%E_{critical} = 1.50\%$)				
	Design	2.43	>	=
	Columns*	1.85	=	=

(Contnd.)

Table 5. Continued

Response	Variance was considered within	$F=\frac{s_Y^2}{s_n^2}$	Conclusion	
			$\alpha=0.05$	$\alpha=0.01$
M_z				
Low levels ($\%E_{critical}=4.95\%$)				
	Design	5.07	>	>
	Columns*	3.54	>	>
	Columns + injection temperature	3.25	>	>
	Columns + injection temperature + analyst	2.68	>	>
High levels ($\%E_{critical}=2.27\%$)				
	Design	1.81	=	=
M_w/M_n				
Low levels ($\%E_{critical}=3.15\%$)				
	Design	5.03	>	>
	Columns**	2.25	>	=
	Columns + detector*	1.02	=	=
High levels ($\%E_{critical}=3.76\%$)				
	Design	3.64	>	>
	Columns	3.28	>	>
	Columns + detector	1.91	=	=
M_z/M_w				
Low levels ($\%E_{critical}=3.24\%$)				
	Design	2.87	>	>
	Columns	2.36	>	=
High levels ($\%E_{critical}=1.46\%$)				
	Design	1.71	=	=

** and *: factors with statistically significant effect at 1% and 5% significance level respectively.
[a] Means H_1: $s_Y^2 > s_n^2$ accepted.
[b] Means H_0: $s_Y^2 = s_n^2$ accepted.

was repeated for the second and the third most important factors, i.e., standardisation for two or three factors.

From Table 5 it can be observed that standardising for the statistically significant factors (indicated with ** and *) reduces the variance of the design experiments, i.e., decreases the F-value to such an extent that s_Y^2 is no longer significantly larger than s_n^2. For the cases where after standardisation for the significant factors H_1: $s_Y^2 > s_n^2$ is still accepted, the following conclusions can be drawn. Standardising non significant factors does not much reduce the s_Y^2 (see M_{top}, M_w and M_z). The fact that after standardising for the significant factors the s_Y^2 is still larger than s_n^2 is not due to the variance caused by the factors themselves but to a larger experimental error within the design. This can be concluded from the above and from the fact that for these responses the $\%E_{critical}$, which represents the experimental error in the design, is larger. The latter can be seen with M_{top}, M_w and M_z (Table 5) when comparing the low levels design with the high levels one. After standardisation on columns there is no reason why the variance s_n^2 (or F) should be larger in the low levels design unless the experimental error is higher, which indeed is the case.

In summary it can be stated that, when the method is standardised for column and detector and the other factors vary within the levels examined in the design, the results obtained would not be relevantly different, from a chromatographic point of view, from the nominal results.

To study the ruggedness of the chromatographic resolution the following approach was taken. Conventionally, resolution between two peaks in a chromatogram is defined as the ratio of the distance

between the peak maxima and the average peak width, viz.

$$R_s = \frac{\Delta t}{1/2(w_1 + w_2)} = \frac{1.18\Delta t}{(w_{1/2,1} + w_{1/2,2})} \quad (10)$$

In SEC a more-meaningful parameter to describe the quality of the separation system (columns plus instrument) is the specific resolution, defined as

$$R_{sp} = \frac{1.18\Delta t}{(w_{1/2,1} + w_{1/2,2})} \frac{1}{\log M_1/M_2} \quad (11)$$

On the linear part of a SEC calibration curve with slope S this becomes

$$R_{sp} = \frac{-1.18/S}{(w_{1/2,1} + w_{1/2,2})} \quad (12)$$

Since SEC calibration curves (log M vs. t) run downward, S always has a negative value. Two factors are seen to affect the resolution: (i) the discriminating power of the columns, which is determined by the total volume and the size distribution of the pores in the particles and (ii) the width of the chromatographic peaks, which is mainly determined by the efficiency of the columns and the quality of the instrumentation.

When a component of a unique molecular weight is injected, it is sensible to express the quality of the set of columns in terms of the equivalent number of theoretical plates. In our case, we used elemental sulphur as an internal standard (flow-rate marker). Plate counts (N) obtained for sulphur on the three different sets of columns (average of duplicate measurements) were 25 200 (PolymerLabs), 28 100 (Phenomenex), and 33 100 (Waters). Since R_s is proportional to $\sqrt{N}$, the effect of the column efficiency on the resolution was +6% for the Phenomenex columns and +15% for the Waters columns, both relative to the PolymerLabs columns.

As seen in Table 6, the effect of the plate count on resolution is overshadowed by that of the size selectivity of the columns. Despite having the lowest plate count for sulphur, the PolymerLabs columns showed the highest specific resolution, especially in the high-MM region. The variation of the specific resolution with variations in the operating conditions during the ruggedness test turned out to be insignificant (typically <0.1 units).

5. Conclusions

The statistical strategy for the identification of significant effects where two-factor interaction effects are used to estimate the experimental error in a fractional factorial design applied in a ruggedness test, are found to be effective for this SEC assay.

The SEC method that was tested here is not so unrugged as was feared before the experiments were started. The factor influencing almost all responses is the column type. This has to be taken into account when columns of different manufacturers are applied. Different columns can have different adsorption characteristics, pore geometries and band broadening which could cause the effects observed in this study.

Another factor is the detector type. For a size-exclusion chromatographer the above conclusions could have been expected but this study allowed to quantify the effects. For the rest it can be concluded that in general a flow-rate of the mobile phase between 0.7 and 0.9 ml/min; a column temperature between 22°C and 32°C; an injection volume between 50 μl and 70 μl; an injection concentration between 2.0 mg/ml and 3.0 mg/ml; an injection

Table 6
Specific-resolution (R_{sp}) values for four different pairs of polystyrene standards on the three different columnsets used in this study

Column	M_1/M_2			
	380 000/96 000	96 000/22 000	22 000/5050	5050/1320
PolymerLabs	4.5	5.6	3.7	2.6
Phenomenex	3.0	4.8	3.8	2.6
Waters	4.0	5.0	3.7	2.6

Data are averages of duplicate measurements, which never differed by more than 0.2 units.

temperature between 22°C and 32°C and a detector cell temperature between 30°C and 40°C do not influence the responses of this method significantly and that variations of these factors do not cause chromatographically relevant changes.

Acknowledgments

The authors thank the National Fund for Scientific Research (NFWO) for financial support. They also thank Mrs. K. Decq and Mrs. A. De Schrijver for the lay-out of some of the figures. Part of this work was funded by the ISC programme of the European Commission under the project CII*-CT93-0097.

Appendix A

A two-level full factorial design is a design where all combinations of factor levels are made. To examine eight factors using such a design requires $2^8 = 256$ experiments. This would allow the determination of all main effects as well as all interaction effects, i.e., a total of 255 effects. The main effects (E_X) are the effects caused by the factors examined in the design. A two-factor interaction effect ($E_{X_iX_j}$) occurs when the main effect of one factor (X_i) is different at both levels of the second one (X_j). A three factor interaction effect ($E_{X_iX_jX_k}$) occurs when the two-factor interaction effects are different at both levels of the third one. Similarly four-factor, five-factor,... interactions can be considered. The physical meaning of these interactions is not evident. For practical reasons it is not possible to perform 256 experiments. Therefore only a fraction of the full factorial design is executed. In our case it was a $1/16^{th}$ fraction, represented as 2^{8-4}, which requires 16 experiments. Reducing the number of experiments implies that some information is lost. The 2^{8-4} design allows to calculate only 15 effects. This means that a number of effects are estimated together. In our case each effect that is calculated is in fact a combination of 16 effects. Such effects are called confounded. The design performed in this study is shown in Table 7. The generators [5] of the design were E = ABC, F = BCD, G = ABD and H = ACD. These generators allow to determine the so-called defining relations, I, [5] of the design which were here equal to I = ABCE = BCDF = ABDG = ACDH = ADEF = CDEG = BDEH = ACFG = ABFH = BCGH = BEFG = CEFH = AEGH = DFGH = ABCDEFGH. The defining relations allow determination of all confoundings in the fractional factorial design [5]. For example the main effect of factor A is confounded with the interactions BCE, CDH, BDG, DEF, CFG, BFH, EGH, ABCDF, ACDEG, ABDEH, ABCGH, ABEFG, ACEFH, ADFGH and BCDEFGH. The interaction effect AB is confounded with the interactions CE, DG, FH, ACDF, BCDH, BDEF, ABCDEG, ADEH, BCFG, ACGH, AEFG, ABCEFH, BEGH, ABDFGH and CDEFGH. It can be seen that the main effect of A is not confounded with two-factor interactions. The same is the case for all other main effects. This is because the generators were chosen so that a design of resolution IV [5] was created. The effect calculated for A, although a combination of 16 effects, is still considered an estimate for the effect of factor A, since three-factor and higher order interaction effects are usually negligible. Main effects are usually larger than two-factor interactions, two-factor interactions are larger than three-factor interactions, etc. Since two-factor

Table 7
Sixteenth-fraction factorial design of resolution IV, 2^{8-4}(IV), with generators E = ABC, F = BCD, G = ABD and H = ACD

	Factors							
Experiment	A	B	C	D	E	F	G	H
1	−	−	−	−	−	−	−	−
2	+	−	−	−	+	−	+	+
3	−	+	−	−	+	+	+	−
4	+	+	−	−	−	+	−	+
5	−	−	+	−	+	+	−	+
6	+	−	+	−	−	+	+	−
7	−	+	+	−	−	−	+	+
8	+	+	+	−	+	−	−	−
9	−	−	−	+	−	+	+	+
10	+	−	−	+	+	+	−	−
11	−	+	−	+	+	−	−	+
12	+	+	−	+	−	−	+	−
13	−	−	+	+	+	−	+	−
14	+	−	+	+	−	−	−	+
15	−	+	+	+	−	+	−	−
16	+	+	+	+	+	+	+	+

interactions are the largest interaction effects we avoided confounding with main effects. In our strategy for ruggedness testing, two-factor interaction effects (which are a confounding of four two-factor interactions and 12 higher order interactions) are also considered negligible. They may therefore be used to estimate the experimental error in a design (see Section 2). As already stated in the theoretical part, the normal probability plots are a tool to verify the negligibility of the two-factor interactions. When the factors are examined at three levels (coded (0)= nominal, (−1) and (+1)=extreme levels) the design is reflected. This means that the design is performed twice, a first time examining levels (0) and (−1), and a second time with (0) and (+1). The design shown in Table 8 is the reflected of Table 7. When the chromatograms resulting from the design of Table 7 are interpreted by two analysts a new design containing nine factors can be created (see Table 9). The effects of factors A till H calculated from this latter design are equal to the average effects calculated by both analysts for these factors. The effect of the ninth factor (T) represents the difference in interpretation by both analysts.

Table 8
Reflected 1/16th fraction factorial design derived from the design of Table 7

	Factors							
Experiment	A	B	C	D	E	F	G	H
1	−1	−1	−1	−1	−1	−1	−1	−1
2	0	−1	−1	−1	0	−1	0	0
3	−1	0	−1	−1	0	0	0	−1
4	0	0	−1	−1	−1	0	−1	0
5	−1	−1	0	−1	0	0	−1	0
6	0	−1	0	−1	−1	0	0	−1
7	−1	0	0	−1	−1	−1	0	0
8	0	0	0	−1	0	−1	−1	−1
9	−1	−1	−1	0	−1	0	0	0
10	0	−1	−1	0	0	0	−1	−1
11	−1	0	−1	0	0	−1	−1	0
12	0	0	−1	0	−1	−1	0	−1
13	−1	−1	0	0	0	−1	0	−1
14	0	−1	0	0	−1	−1	−1	0
15	−1	0	0	0	−1	0	−1	−1
16	0	0	0	0	0	0	0	0
17	+1	+1	+1	+1	+1	+1	+1	+1
18	0	+1	+1	+1	0	+1	0	0
19	+1	0	+1	+1	0	0	0	+1
20	0	0	+1	+1	+1	0	+1	0
21	+1	+1	0	+1	0	0	+1	0
22	0	+1	0	+1	+1	0	0	+1
23	+1	0	0	+1	+1	+1	0	0
24	0	0	0	+1	0	+1	+1	+1
25	+1	+1	+1	0	+1	0	0	0
26	0	+1	+1	0	0	0	+1	+1
27	+1	0	+1	0	0	+1	+1	0
28	0	0	+1	0	+1	+1	0	+1
29	+1	+1	0	0	0	+1	0	+1
30	0	+1	0	0	+1	+1	+1	0
31	+1	0	0	0	+1	0	+1	+1
32	0	0	0	0	0	0	0	0

Table 9
Combined design created from Table 7 by interpretation of the design results by two analysts

	Factors								
Experiment	A	B	C	D	E	F	G	H	T
1	−	−	−	−	−	−	−	−	+
2	+	−	−	−	+	−	+	+	+
3	−	+	−	−	+	+	+	−	+
4	+	+	−	−	−	+	−	+	+
5	−	−	+	−	+	+	−	+	+
6	+	−	+	−	−	+	+	−	+
7	−	+	+	−	−	−	+	+	+
8	+	+	+	−	+	−	−	−	+
9	−	−	−	+	−	+	+	+	+
10	+	−	−	+	+	+	−	−	+
11	−	+	−	+	+	−	−	+	+
12	+	+	−	+	−	−	+	−	+
13	−	−	+	+	+	−	+	−	+
14	+	−	+	+	−	−	−	+	+
15	−	+	+	+	−	+	−	−	+
16	+	+	+	+	+	+	+	+	+
1'	−	−	−	−	−	−	−	−	−
2'	+	−	−	−	+	−	+	+	−
3'	−	+	−	−	+	+	+	−	−
4'	+	+	−	−	−	+	−	+	−
5'	−	−	+	−	+	+	−	+	−
6'	+	−	+	−	−	+	+	−	−
7'	−	+	+	−	−	−	+	+	−
8'	+	+	+	−	+	−	−	−	−
9'	−	−	−	+	−	+	+	+	−
10'	+	−	−	+	+	+	−	−	−
11'	−	+	−	+	+	−	−	+	−
12'	+	+	−	+	−	−	+	−	−
13'	−	−	+	+	+	−	+	−	−
14'	+	−	+	+	−	−	−	+	−
15'	−	+	+	+	−	+	−	−	−
16'	+	+	+	+	+	+	+	+	−

The ninth factor (T) represents the two analysts (level (+)=T1, level (−)=T2).

References

[1] J.A. Van Leeuwen, L.M.C. Buydens, B.G.M. Vandeginste, G. Kateman, P.J. Schoenmakers and M. Mulholland, Chemometrics and Intelligent Laboratory Systems, 11 (1991) 37–55.
[2] Y. Vander Heyden, K. Luypaert, C. Hartmann, D.L. Massart, J. Hoogmartens and J. De Beer, Anal. Chim. Acta, 312 (1995) 245–262.
[3] Y. Vander Heyden, C. Hartmann, D.L. Massart, L. Michel, P. Kiechle and F. Erni, Anal. Chim. Acta, 316 (1995) 15–26.
[4] J.L. Virlichie and A. Ayache, S.T.P. Pharma Pratiques, 5 (1995) 49–60.
[5] E. Morgan, Chemometrics: Experimental Design; Analytical Chemistry by Open Learning, Wiley, Chichester, 1991, pp. 110–113, pp. 161–171.
[6] G. Box, W. Hunter and J. Hunter, Statistics for Experimenters, an Introduction to Design, Data analysis and Model Building, Wiley, New York, 1978, pp. 306–418.
[7] Y. Vander Heyden, D.L. Massart, Y. Zhu, J. Hoogmartens and J. De Beer, J. Pharm. Biomed. Anal., 14 (1996) 1313–1326.
[8] G.T. Wernimont, in W. Spendley (Editor), Use of Statistics to Develop and Evaluate Analytical Methods, Association of Official Analytical Chemists, Arlington, VA, USA, 1985, pp. 78–82.
[9] H. Fabre, V. Meynier de Salinelles, G. Cassanas and B. Mandrou, Analusis, 13 (1985) 117–123.
[10] K. Jones, Int. Lab., November (1986) 32–45.
[11] S.F.Y. Li, Capillary Electrophoresis: Principles, Practice and Applications, Journal of Chromatography Library, Vol. 52, Elsevier, Amsterdam, 1992, pp. 316–318.
[12] T. Welsch, 19th International Symposium on Column Liquid Chromatography, Innsbruck, Austria, May 28th–June 2nd, 1995.
[13] S. Boonkerd, M.R. Detaevernier, Y. Vander Heyden, J. Vindevogel and Y. Michotte, J. Chromatogr. A, 736 (1996) 281–289.
[14] Ø. Langsrud, M.R. Ellekjaer and T. Naes, J. Chemometrics, 8 (1994) 205–219.
[15] G.M.R. Vandenbossche, C. De Muynck, Y. Vander Heyden, D.L. Massart, K. Strobbe, P. Van Aerde and J.P. Remon, personal communication.
[16] C. Daniel, Applications of Statistics to Industrial Experimentation, Wiley, New-York, 1976, 274 pp.
[17] J. Goupy, Chemometrics and Intelligent Laboratory systems, in press.
[18] G. Box and R.D. Meyer, J. Res. Natl. Bureau of Standards, 90 (1985) 495–500.