Journal of Chromatography, 522 (1990) 163–170
Elsevier Science Publishers B.V., Amsterdam

CHROM. 22 761

Reversed-phase chromatography of *Escherichia coli* ribosomal proteins

Correlation of retention time with chain length and hydrophobicity

W. SCOTT CHAMPNEY
Department of Biochemistry, College of Medicine, East Tennessee State University, Johnson City, TN 37614 (U.S.A.)

(First received February 28th, 1990; revised manuscript received July 24th, 1990)

ABSTRACT

The elution times of 52 bacterial ribosomal proteins from a C_4 reversed-phase column have been predicted. The prediction is based upon the use of the hydrophobicity coefficients for the protein amino acid content as defined by Guo *et al.* [*J. Chromatogr.*, 359 (1986) 499–518]. A strong positive correlation was observed when the difference between predicted and observed protein retention time was plotted against the product of net hydrophobicity and natural log of protein chain length. Observations with this class of related proteins strengthens and extends the observations of Mant *et al.* [*J. Chromatogr.*, 476 (1989) 363–375]. Observed deviations from predicted chromatographic behavior can be explained for several proteins which elute as dimers or which have modified amino acid residues.

INTRODUCTION

Reversed-phase high-performance liquid chromatography has been demonstrated by several groups to be an effective and efficient method for the separation of the complex mixture of proteins found as a part of the bacterial ribosome structure [1–3]. Proteins from both subunits of the bacterial ribosome have been separated and identified by this method, using a variety of similar chromatographic techniques. Both analytical and preparative separations have been performed. The analysis has recently been extended to examine questions related to the function of particular proteins in the process of translation [4,5].

The sequences of all of these proteins are known [6]. Information about the amino acid composition has permitted a calculation of the coefficient of hydrophobicity for each protein, using the values derived by Guo *et al.* [7]. Protein hydrophobicity and chain length can be used to accurately predict the retention time for a number of unrelated proteins under several conditions of reversed-phase chromatography [8].

In this work we show the application of this method to the elution behavior of 52 bacterial ribosomal proteins, as mixtures derived from the two ribosomal subunits.

0021-9673/90/$03.50

A very strong correlation between predicted and expected elution time was found. In addition, certain predictions about protein structure and protein-protein interactions were confirmed.

EXPERIMENTAL

Materials

Acetonitrile and trifluoroacetic acid (TFA) were purchased from Pierce, as were the C_4 cartridge (BU-300, 100 × 4.6 mm I.D.) and guard columns. Ribosomal proteins were prepared from the subunits of *Escherichia coli* strain SK901 as previously described [9]. The dried proteins were dissolved in 100 μl of 66% aqueous acetic acid prior to injection.

Reversed-phase high-performance liquid chromatography

The Waters HPLC system consisted of a pair of Model 510 pumps, a U6K injector, a Model 680 gradient controller, a Model 481 Lambda Max detector and a Model 730 data module. A C_4 (butyl) cartridge and guard column were used for all separations. Gradients were run at 0.25 ml/min using 0.1% TFA in Milli-Q water (pH 2.0) as the stationary phase and 0.1% TFA in acetonitrile as the mobile phase. The detector was set at 215 nm. Samples of 400 μg of 30S proteins and 600 μg of 50S proteins were injected in 100 μl of 66% acetic acid. The proteins were eluted with a linear gradient of 30% to 50% acetonitrile for 155 min followed by a second linear gradient of 50% to 60% for 30 min. Retention times and peak areas were determined by the data module for each run. Column fractions (0.5 ml) were freeze-dried and examined for protein content by one and two dimensional gel electrophoresis as described previously [9–11].

RESULTS AND DISCUSSION

As others have shown previously, reversed-phase HPLC is an excellent method for resolving the complex mixture of similar proteins from the bacterial ribosome [1–3]. Fig. 1 shows the separation of the 30S subunit proteins of *Escherichia coli* on a C_4 column using a linear gradient for elution. The 21 proteins were resolved into 17 peaks. Identification of the proteins was carried out by one- and two-dimensional gel electrophoresis of the column fractions. The separation of the 50S subunit proteins under the same conditions is shown in Fig. 2. The 32 proteins in this mixture were resolved into 25 peaks as shown.

It has been demonstrated by Mant *et al.* [8] that the elution of different proteins from reversed-phase columns can be predicted based on the total hydrophobicity and chain length of the protein. They derived the following relationship: $\Sigma R_c - t_R = A(\Sigma R_c \ln N) + C$; where ΣR_c = the summed hydrophobicity coefficients of the protein amino acids (after Guo and co-workers [7,12]), t_R = the observed protein retention time and $\ln N$ = the natural log of the number of amino acids in the protein. A plot of this linear relationship gives a line with a slope of A and a y-intercept of C.

We have applied this method to predict the elution times of the individual 30S and 50S subunit proteins from the C_4 column. Using the data of Giri *et al.* [6] for the

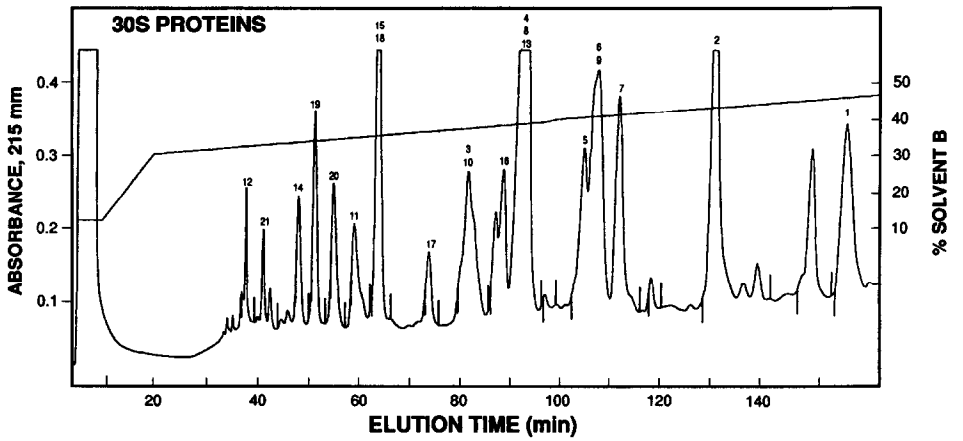

Fig. 1. C_4 reversed-phase HPLC elution profile of 30S subunit proteins. The separation of 400 μg of proteins with a linear gradient of acetonitrile (solvent B) was conducted as described in Experimental. The protein identification and gradient slope are indicated on the tracing.

amino acid composition and chain length of each protein, we have calculated the total hydrophobicity for each protein. Minor corrections in the amino acid composition of proteins S4, S9, S15, S18, L9 and L10 were made based on the DNA sequences of the genes [13–17]. The relevant values for the 30S subunit ribosomal proteins are shown in Table I. The 50S subunit protein values are compiled in Table II.

Mant *et al.* [8] have shown that the difference between predicted (ΣR_c) and observed (t_R) retention times for 23 distinct proteins is a linear function of $\Sigma R_c \cdot \ln N$, the protein hydrophobicity and chain length. This strong correlation is also true for the separated ribosomal proteins as Figs. 3 and 4 indicate. A correlation of 0.996 for the predicted and expected elution of the 30S subunit proteins was found. For the 50S proteins the correlation was 0.964. For these 52 related proteins the product $\Sigma R_c \cdot \ln N$ was a very accurate predictor of retention time.

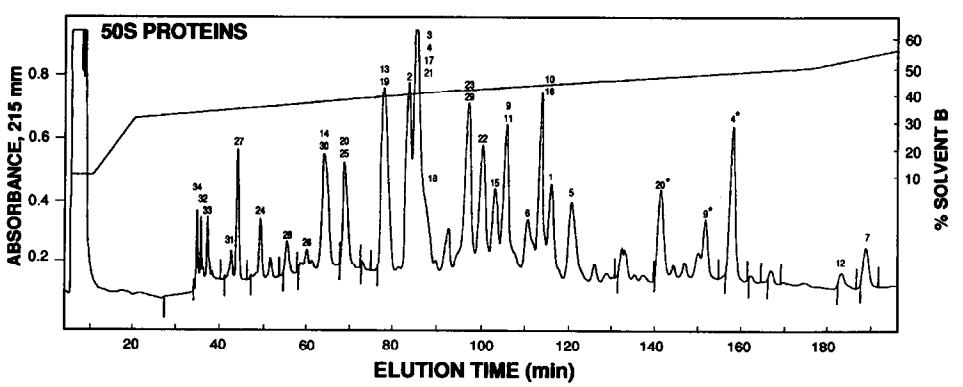

Fig. 2. C_4 reversed-phase HPLC elution profile of 50S subunit proteins. The separation of 600 μg of proteins with a linear gradient of acetonitrile (solvent B) was conducted as described in Experimental. The protein identification and gradient slope are indicated on the tracing. 4*, 9* and 20* indicate the second, later eluting positions for these proteins.

TABLE I

30S SUBUNIT PROTEIN HYDROPHOBICITY COEFFICIENTS AND RETENTION TIMES

Protein	N^a	$\Sigma R_c{}^b$	$\Sigma R_c \cdot \ln N$	$t_R{}^c$	$\Sigma R_c - t_R{}^d$
S1	557	1230.4	7779.3	158.0(0.51)	1072.4
S2	240	544.4	2983.7	129.4(0.65)	415.0
S3	232	471.4	2567.6	92.5(0.66)	378.9
S4	205	398.3	2120.2	93.0(0.66)	305.3
S5	166	331.0	1692.1	105.9(0.51)	225.1
S6	135	261.9	1284.7	106.1(0.34)	155.8
S7	177	358.7	1856.7	111.4(0.76)	247.3
S8	129	285.1	1385.5	93.0(0.66)	192.1
S9	129	238.1	1157.1	104.2(0.54)	133.9
S10	103	231.2	1071.5	87.3(0.82)	143.9
S11	128	217.3	1054.4	59.7(0.55)	157.6
S12	123	183.6	883.5	38.6(0.42)	145.0
S13	117	218.1	1038.6	92.5(0.66)	125.6
S14	98	161.5	740.5	48.8(0.29)	112.7
S15	88	158.4	709.2	63.5(0.62)	94.9
S16	82	169.3	746.1	80.7(0.88)	88.6
S17	83	170.6	753.9	73.0(0.57)	97.6
S18	74	156.7	675.3	63.5(0.62)	93.2
S19	91	157.7	711.4	52.2(0.40)	105.5
S20	86	115.8	515.8	56.2(0.39)	59.6
S21	70	101.6	431.7	41.8(0.35)	59.8

[a] Number of amino acid residues from Giri *et al.* [6].
[b] Sum of retention coefficients calculated from values of Guo *et al.* [7].
[c] Mean of observed retention times ($n = 6$) with standard error in parentheses.
[d] Difference between predicted and observed retention times.

Post translational modifications in the form of methyl and acetyl group additions are found for several of the ribosomal proteins [6]. Corrections for the presence of these groups and their influence on the ΣR_c for the affected proteins have been calculated (Table III). The presence of a methyl group was predicted to increase the ΣR_c by $+1.8$, the difference in R_c between glycine and alanine [7]. For example, protein L11 contains 9 methyl groups. The difference in predicted retention times for the unmodified and modified forms of L11 are indicated in Table III and in Fig. 4. The method is sensitive enough to reveal differences in elution time as a consequence of these modifications. Corrections have also been made for the N-terminal acetylation of proteins S5, S18 and L7 which increases the ΣR_c by a value of $+6.9$ [7]. Proteins S12 and L16 each contain an additional modified amino acid [6]. Corrections have not been made for these changes since the nature of the alteration and the contribution to the protein hydrophobicity are not known.

Ribosomal protein L7 is the N-terminal acetylated form of protein L12. This alteration is responsible for the slight difference in retention time for these two proteins. Both of these proteins exist in the ribosomal particle as dimers, unlike any of the other ribosomal proteins [18]. It is obvious from the observed retention times for these two proteins that they eluted from the C_4 column as dimers and not in the position expected for the monomeric form of each protein (Fig. 4). For the L7 dimer

TABLE II

50S RIBOSOMAL PROTEIN HYDROPHOBICITY COEFFICIENTS AND RETENTION TIMES

Protein	N^a	$\Sigma R_c{}^b$	$\Sigma R_c \cdot \ln N$	$t_R{}^c$	$\Sigma R_c - t_R{}^d$
L1	233	466.1	2540.7	113.0(0.91)	353.1
L2	272	426.8	2392.6	78.8(0.62)	348.0
L3	209	400.7	2140.5	81.4(0.66)	319.3
L4	201	432.0	2291.0	81.4(0.66)	350.6
L5	178	421.1	2182.0	117.8(1.05)	303.3
L6	176	341.9	1767.8	108.6(1.2)	233.3
L7	120	272.2	1303.0	187.0(0.98)	85.2
L7(d)[e]	240	272.2	2981.0	187.0(0.98)	352.0
L9	148	334.6	1675.3	93.2(0.55)	241.6
L10	165	377.2	1926.2	109.9(0.91)	267.8
L11	141	299.2	1480.7	104.4(0.93)	194.8
L12	120	265.3	1270.1	181.3(0.89)	84.0
L12(d)[e]	240	265.3	2906.0	181.3(0.89)	346.0
L13	142	286.4	1419.4	73.0(0.65)	213.4
L14	123	266.9	1284.4	66.3(1.0)	200.6
L15	144	263.9	1311.5	104.4(0.93)	159.5
L16	136	290.1	1425.2	110.6(1.16)	179.5
L17	127	245.3	1188.3	81.4(0.66)	163.9
L18	117	210.3	1001.5	81.4(0.66)	128.9
L19	114	215.6	1021.1	73.0(0.66)	142.6
L20	117	238.9	1144.3	70.1(1.58)	170.2
L21	103	199.4	924.2	81.4(0.66)	118.0
L22	110	201.9	949.0	99.0(0.77)	102.9
L23	99	183.4	842.7	96.6(0.77)	86.8
L24	103	176.9	819.9	46.9(1.2)	130.0
L25	94	190.7	866.4	66.3(1.0)	124.4
L26[f]	86	115.8	515.8	56.0(0.39)	59.8
L27	84	107.6	476.8	42.9(0.27)	64.7
L28	77	131.2	569.9	54.6(0.84)	76.6
L29	63	128.0	530.3	93.2(0.55)	34.8
L30	58	118.2	479.9	60.3(0.67)	57.9
L31	62	92.5	381.8	42.8(0.26)	49.7
L32	56	55.0	221.4	35.8(0.19)	19.2
L33	54	78.0	311.1	36.7(0.24)	41.3
L34	46	50.4	192.9	34.7(0.17)	15.7

[a] Number of amino acid residues from Giri *et al.* [6].
[b] Sum of retention coefficients calculated from values of Guo *et al.* [7].
[c] Mean of observed retention times ($n = 9$) with standard error in parentheses.
[d] Difference between predicted and observed retention times.
[e] Proteins L7 and L12 as dimers (d).
[f] Protein L26 is the same as protein S20 [6].

the product $\Sigma R_c \cdot \ln N$ increases from 1302 to 2981 and for the L12 dimer the value changes from 1270 to 2906. Protein L7 recovered after the HPLC separation was found to elute from a calibrated gel filtration column at the position expected for a protein of 24 000 dalton, the size of the L7 dimer (unpublished observations).

A similiar explanation may apply to the observed elution of 50S proteins L4, L9 and L20. These three proteins have each been found to elute in two distinct locations

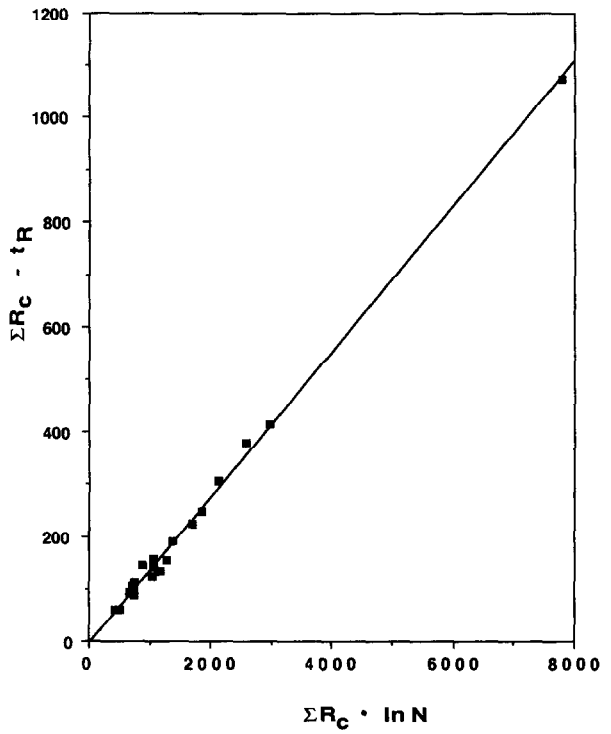

Fig. 3. Correlation of 30S protein retention times with polypeptide chain length and hydrophobicity. Predicted *minus* observed retention time $\Sigma R_c - t_R$ *versus* $\Sigma R_c \cdot \ln N$. The slope of this line is defined by $y = 0.1389\ x - 2.766$ ($r = 0.996$).

from the column, as Fig. 2 indicates. The later elution position of each is consistent with the expected location of each as a protein dimer. There is no indication that any of these function as dimers in the 50S subunit [6]. Alternatively, the later elution positions may reflect interactions between these and other proteins which are only disrupted by the increased acetonitrile concentration during the chromatography. The anomalous elution of protein L9 has been noted previously [4].

Finally, we have investigated the elution profiles of 50S proteins from a pair of temperature-sensitive mutants of *E. coli* with independent alterations in protein L22 [9]. Both mutant proteins are less basic than the normal protein as revealed by their mobility in two-dimensional gel electrophoresis. Both altered proteins eluted as less hydrophobic species under these chromatography conditions, with retention times decreased by 1 and 2 min, respectively. A single amino acid change in each protein was sufficient to promote this difference in elution time, indicating the specificity and sensitivity of this separation method.

The excellent agreement found between the expected and predicted elution behavior of these 52 proteins supports the suggestions of Mant *et al.* [8] that protein hydrophobicity and chain length are the primary determinants of chromatographic properties under these conditions. In 0.1% TFA at pH 2, these proteins should be fully denatured and thus allow maximal interaction between the amino acid sequences and the hydrophobic column matrix. The use of a linear gradient under the same

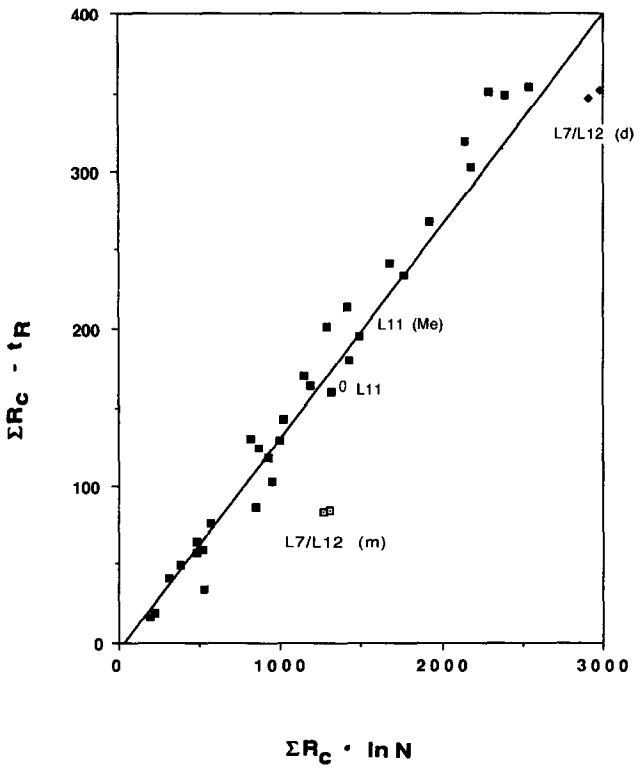

Fig. 4. Correlation of 50S protein retention times with polypeptide chain length and hydrophobicity. Predicted *minus* observed retention time $\Sigma R_c - t_R$ *versus* $\Sigma R_c \cdot \ln N$. The expected locations of the L7/L12 monomer (m) and dimers (d) are indicated as are the positions for the unmodified (0) and methylated (Me) forms of L11. The slope of this line is defined by $y = 0.1368\ x - 3.678$ ($r = 0.964$).

TABLE III

RIBOSOMAL PROTEIN MODIFICATION AND EFFECTS ON ΣR_c

Protein	Modification	ΣR_c	
		Unmodified	Modified[a]
S5	N-Acetyl-Ala	324.1	331.0
S11	N-Me-Ala	208.6	217.3
S18	N-Acetyl-Ala	149.8	156.7
L3	N-Me-Gln	398.9	400.7
L7	N-Acetyl-Ser		
	N-Me-Lys	263.5	272.2
L12	N-Me-Lys	263.5	265.3
L11	$N\text{-}Me_3\text{-}Ala$		
	$(N\text{-}Me_3\text{-}Lys)_2$	283.0	299.2
L16	N-Me-Met	288.3	290.1
L33	N-Me-Ala	76.2	78.0

[a] A value of +1.8 was added for each methyl group (Me) and a value of +6.9 was added for each blocked N-terminal amino acid.

solvent conditions as Guo and co-workers [7,12] and Mant and co-workers [8,19] allowed a direct correlation of these protein hydrophobicity values with their predicted hydrophobicity coefficients.

We have also used this method to compare the elution pattern of the ribosomal proteins published by Kerlavage *et al.* [1], using their α values to compute retention times. Although they used hyperbolic gradients for elution of the proteins from a C_{18} column, a good correlation was found between the observed and expected elution of the 30S proteins under their conditions (data not shown). The relationship for the expected elution of the 50S proteins was not as good as observed in the present work. Other workers [2,3] have used more complex gradient procedures to optimize ribosomal protein separation, making a direct comparison with the present results difficult.

We have extended the work of Mant *et al.* [8] to a complex group of related proteins. We have shown that their rules for predicting protein retention time apply to the ribosomal proteins separated under the conditions described. Their methods also allow predictions about protein–protein interactions (as dimers), about post translational modifications and about the elution of mutationally altered protein molecules. This method should have wide application for predicting the separation of other well characterized protein complexes, for identifying the principles underlying separations on reversed-phase columns and for analyzing modified forms of similiar protein molecules.

ACKNOWLEDGEMENTS

The excellent technical assistance of Kiersten Hutchinson is gratefully acknowledged. This work was supported by NIH grant GM 38492.

REFERENCES

1 A. R. Kerlavage, T. Hasan and B. S. Cooperman, *J. Biol. Chem.*, 258 (1983) 6313–6318.
2 R. M. Kamp, A. Bosserhoff, D. Kamp and B. Wittmann-Liebold, *J. Chromatogr.*, 317 (1984) 181–192.
3 R. J. Ferris, C. A. Cowgill and R. R. Traut, *Biochemistry*, 23 (1984) 3434–3442.
4 B. S. Cooperman, C. J. Weitzmann and M. A. Buck, *Methods Enzymol.*, Vol. 164, 1988, pp. 523–532.
5 A. R. Kerlavage, C. Weitzmann, M. Cannon, T. Hasan, K. M. Giangiacomo, J. Smith and B. S. Cooperman, *BioTechniques*, Jan./Feb. (1985) 26–36.
6 L. Giri, W. E. Hill, H. G. Wittmann and B. Wittmann-Liebold, in C. B. Anfinsen, J. T. Edsall and F. M. Richards (Editors), *Advances in Protein Chemistry*, Academic Press, New York, 1984, pp. 1–78.
7 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499–518.
8 C. T. Mant, N. E. Zhou and R. S. Hodges, *J. Chromatogr.*, 476 (1989) 363–375.
9 W. S. Champney, *Biochim. Biophys. Acta*, 609 (1980) 464–474.
10 G. A. Howard and R. R. Traut, *FEBS Lett.*, 29 (1973) 177–180.
11 D. B. Datta, L.-M. Changchien, C. R. Nierras, W. A. Strycharz and G. R. Craven, *Anal. Biochem.*, 173 (1988) 241–245.
12 D. Guo, C. T. Mant, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 519–532.
13 D. Bedwell, G. Davis, M. Gosink, L. Post, M. Nomura, H. Kestler, J. M. Zengel and L. Lindahl, *Nucl. Acids Res.*, 13 (1985) 3891–3903.
14 S. Isono, S. Thamm, M. Kitakawa and K. Isono, *Molec. Gen. Genet.*, 198 (1985) 279–282.
15 L. E. Post, G. D. Strycharz, M. Nomura, H. Lewis and P. P. Dennis, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 1697–1701.
16 J. Schnier, M. Kitakawa and K. Isono, *Molec. Gen. Genet.*, 204 (1986) 126–132.
17 R. Takata, T. Mukai, M. Aoyagi and K. Hori, *Molec. Gen. Genet.*, 197 (1984) 225–229.
18 W. Moller and J. A. Maassen, in B. Hardesty and G. Kramer (Editors), *Structure, Function and Genetics of Ribosomes*, Springer, New York, 1986, pp. 309–325.
19 C. T. Mant, T. W. L. Burke, J. A. Black and R. S. Hodges, *J. Chromatogr.*, 458 (1988) 193–205.