Analysis of Single Amino Acid Mutations in Intact Proteins

Robert Freeman[1], Dan DiFeo[1], Kayte Parlevliet[2], Hans-Jürgen Wirth[2]

[1]SGE Incorporated, 2007 Kramer Lane, Austin Texas, USA, [2] SGE Analytical Science, 7 Argent Place, Ringwood Victoria, Australia.
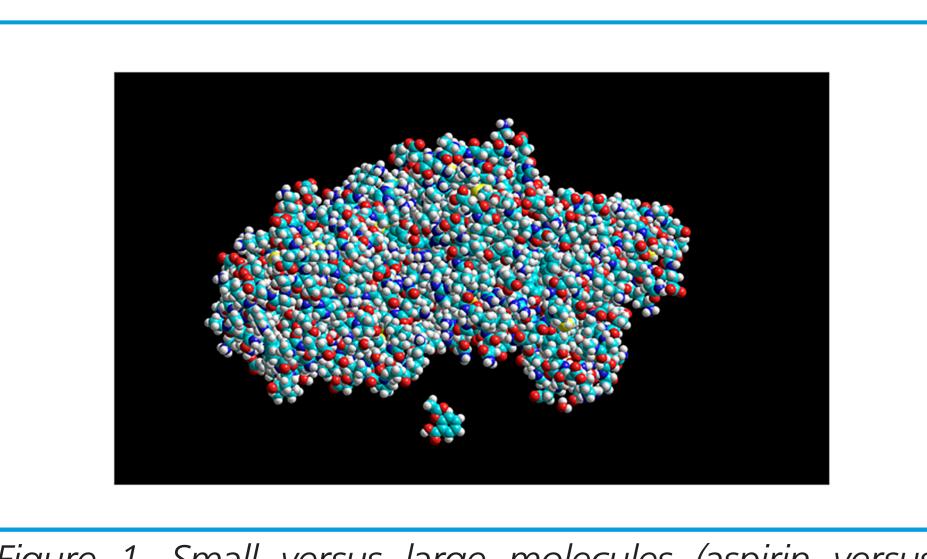
Abstract

With the increasing availability of high resolution mass spectrometers the analysis of intact proteins becomes a feasible alternative to the more commonly employed shot-gun techniques in proteomics. High resolution separations of proteins are hindered by the size of the analyte molecules and the resulting low diffusion rates especially inside the pore system of the stationary phase.

The stationary phase of the columns employed in these studies is a 3 µm C8 silica with 1000 Å pore size. By using a dedicated protein separation column (ProteCol™ C8 HQ1003) we demonstrate that it is possible to separate ribosomal proteins derived from a rat liver with excellent resolution and identify single amino acid mutants in a 33 kDa protein.
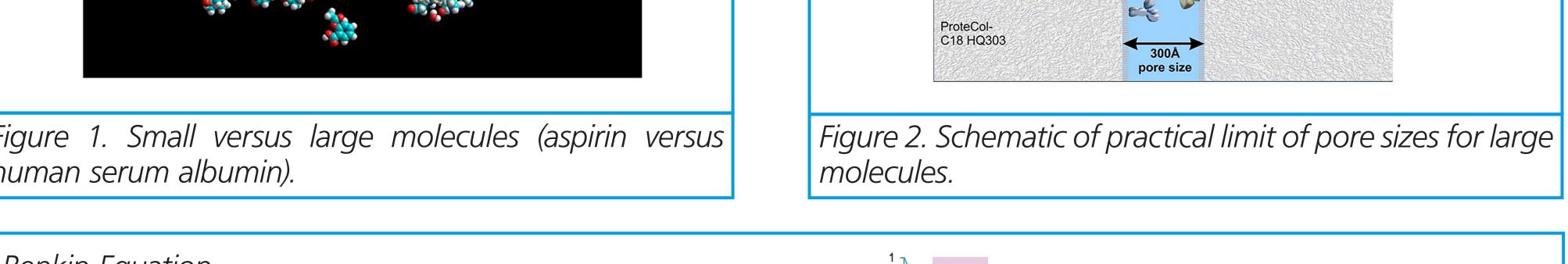
Background

Protein Separation and Diffusion

Diffusion is the time-limited process in the separation of macro-molecules such as proteins. Due to their size, the diffusion rate of proteins in free solution is significantly lower than the diffusion rate of small molecules (see Figure 1). In addition, the effect of restriction inside the pores of the stationary phase is far more pronounced (Figure 2 and 3). To minimize the effects of hindered pore diffusion, a 1000 Å 3 µm C8-silica column (ProteCol™ C8 HQ1003) was employed for this study.
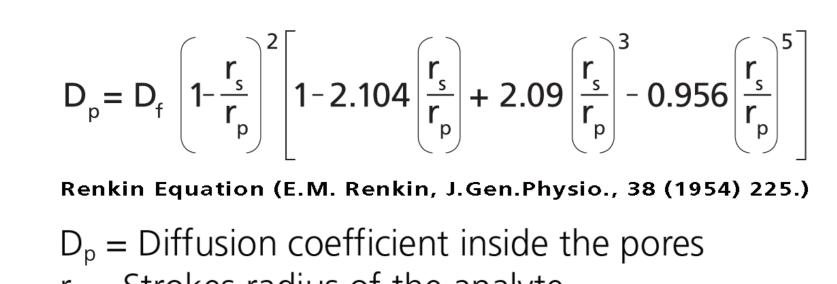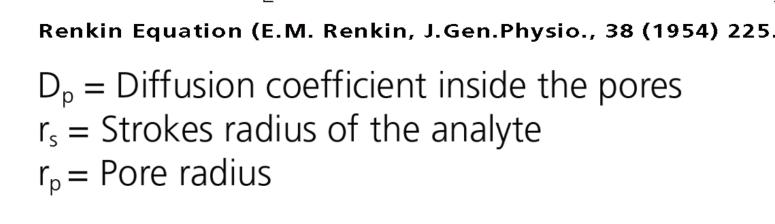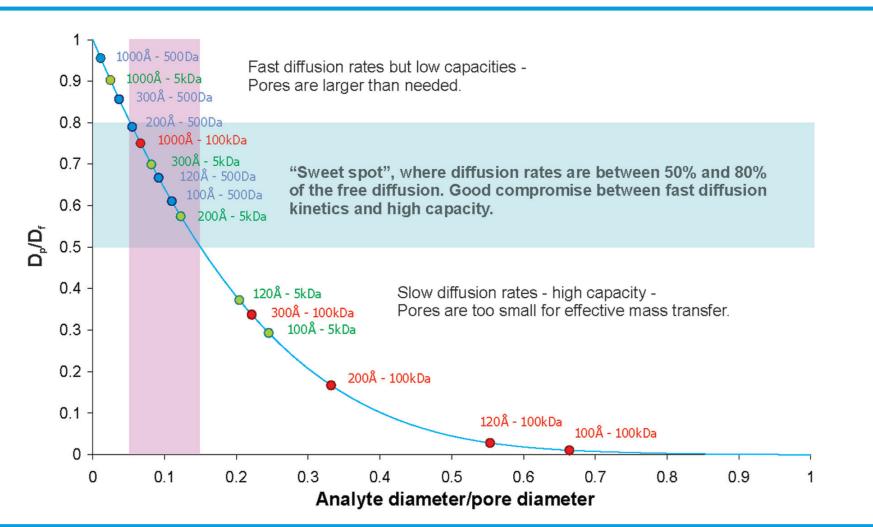

Figure 1. Small versus large molecules (aspirin versus human serum albumin).

Figure 2. Schematic of practical limit of pore sizes for large molecules.

Renkin Equation

$$D_p = D_f \left[1 - \frac{r_s}{r_p} \right]^2 \left[1 - 2.104 \left(\frac{r_s}{r_p} \right) + 2.09 \left(\frac{r_s}{r_p} \right)^3 - 0.956 \left(\frac{r_s}{r_p} \right)^5 \right]$$

Renkin Equation (E.M. Renkin, J.Gen.Physio., 39 (1954) 225.)

D_p = Diffusion coefficient inside the pores
r_s = Strokes radius of the analyte
r_p = Pore radius
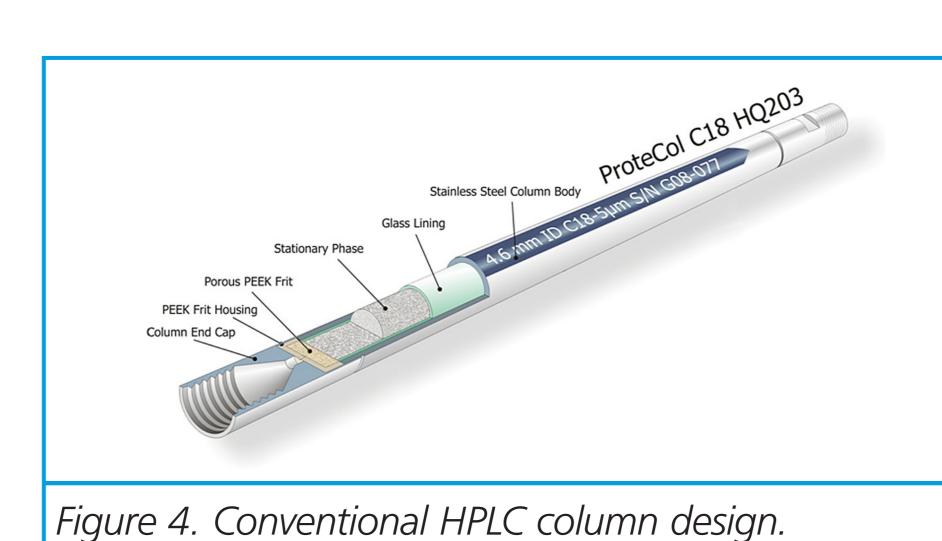

Figure 3. Relative diffusion rates in free solution and inside pores as function of molecule size to pore ratio.

Ribosome Protein Identification

The rat liver 80S ribosome is made up of two subunits – the large subunit has ~49 proteins and the small subunit has ~33 proteins. Since ribosomal proteins are relatively small (Mw=6,000 to 40,000), they can be rapidly identified by accurate LCMS analysis of the intact proteins.

Dedicated Protein Columns – C8, 3 µm, 1000 Å

Figure 4. Conventional HPLC column design.

A glass lined HPLC column of conventional design was used with 0.5 µm pore sized PEEK™ frits within a stainless steel body to ensure an inert flow path for the sample (Figure 4). The C8-silica stationary phase with 3 µm particles with 1000 Å pore size is ideal for the separation of intact proteins.
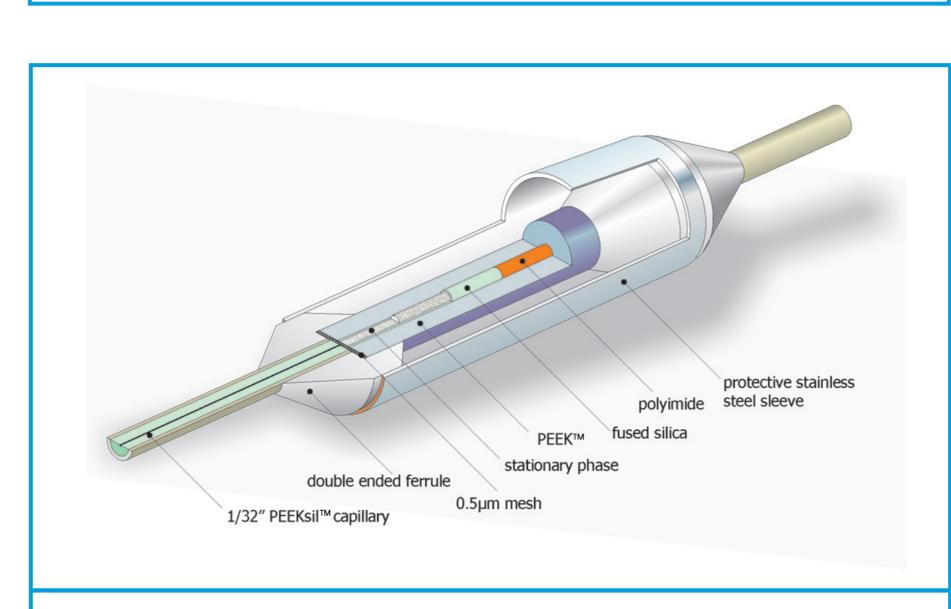

Figure 5. Design of the ProteCol™ range of columns.

The capillary column is made of a packed 1/16" OD PEEKsil™ (polyether ethyl ketone coated fused silica) capillary, with 1/32" OD, 50 µm ID PEEKsil™ connection capillaries permanently attached (Figure 5). The frits are made from a 100 µm thick woven mesh with 0.5 µm pore size. The column is protected by a stainless steel sleeve. Using PEEKsil™ rather than fused silica increases the robustness of the column while maintaining flexibility to easily attach the column to the LC system. The same stationary phase as above was used.

Methodology/Procedure

Analysis of Ribosomal Proteins

Step 1. Sample Preparation

80S ribosomal proteins were isolated from a rat liver microsomal preparation[1]. One optical density unit at 260 nm of 80S ribosomal proteins was mixed with 2 volumes of 6M Guanidine HCl to denature the proteins. 1 % (v/v) formic acid was subsequently added to precipitate the nucleic acids. The mixture was centrifuged for 15 minutes at 13,000 rpm and the supernatant was collected into a sample vial ready for LCMS analysis.

Step 2. LCMS Analysis

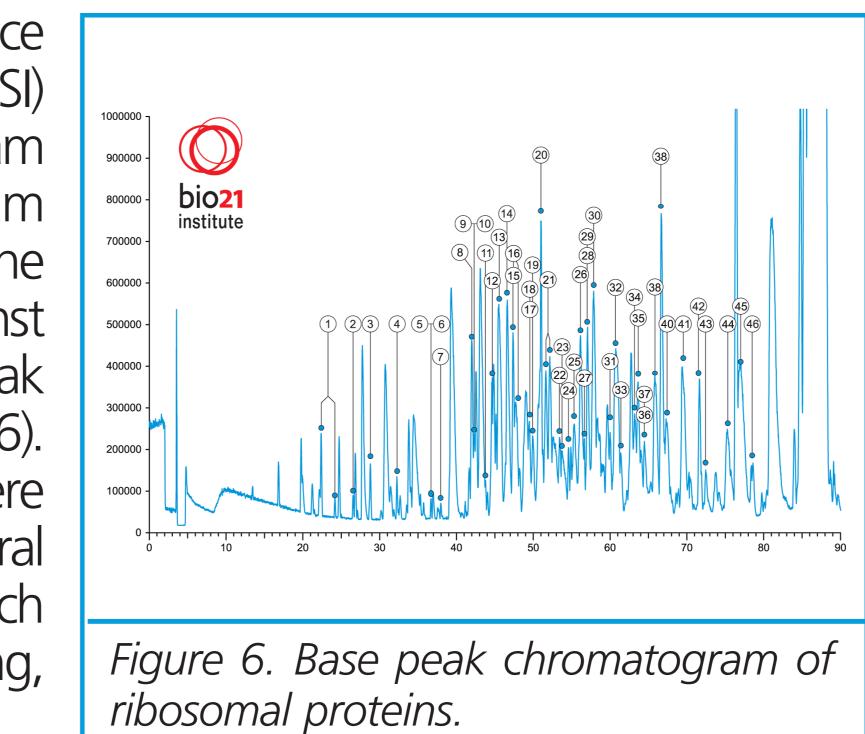|	
System:	Agilent 1100 LC with Agilent 6220 ESI-TOF LC/MS Mass detection
Column:	ProteCol™ – C8 HQ1003 3 µm, 1000 Å 150 mm x 2 mm ID
Flow rate:	0.25 µL/min
Temperature:	80 °C
Solvent A:	Aqueous 0.1 % (v/v) Formic Acid
Solvent B:	Acetonitrile 0.1 % (v/v) Formic Acid
Gradient:	0 min 5 % B 80 min 45 % B 81 min 85 % B
	82 min 85 % B 83 min 5 % B 88 min 5 % B

Analysis of Single Amino Acid Proteins

Nine single-amino acid mutations of a bacterial 33 kDa protein toxin were analysed by RP-µLC-MS in order to confirm the correct structure of the mutant protein. Samples supplied at concentrations ranging from 0.3 to 1 mg/ml. 100 µL of samples were desalted with spin columns and resultant filtrate directly injected at a volume to equal the injection of 1 µg i.e. 2 µL injected of 0.5 mg/mL solution assuming no losses at filtration step.
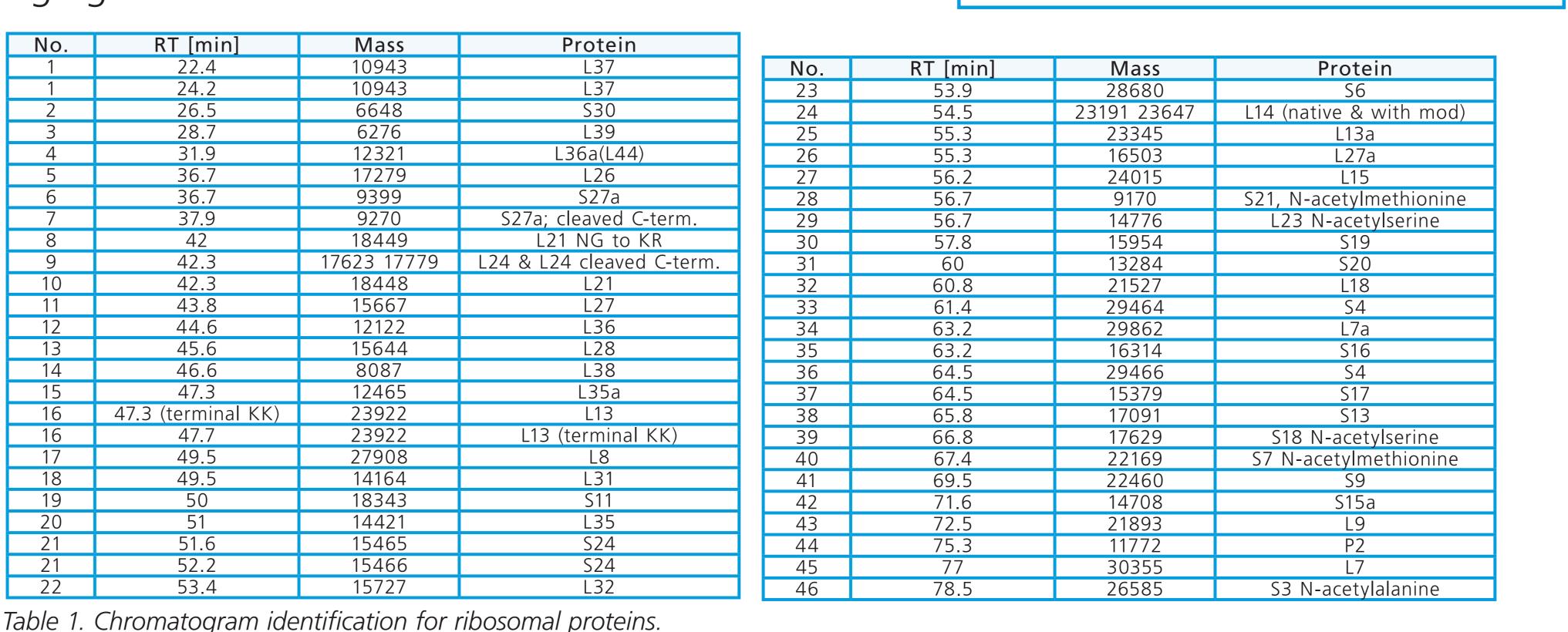
Chromatographic conditions:

Instrument:	Agilent 1200 Cap-LC with Bruker microTOF-Q MS
Column:	ProteCol™ – G C8 HQ1003, 300 µm ID x 100 mm
Flow rate:	5 µL/min
Mobile phase A:	0.1 % formic acid
Mobile phase B:	95 % acetonitrile, 0.1 % formic acid
Gradient:	0 min 5 % B 30 min 55 % B 33 min 70 % B
	38 min 70 % B 40 min 5 % B

Results
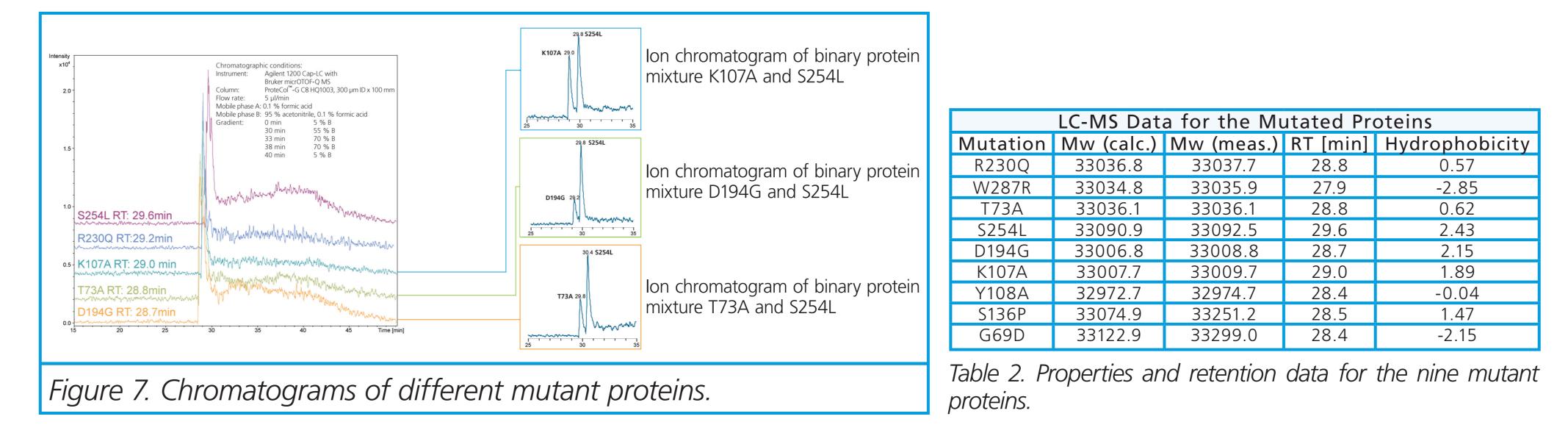
Data for the ribosomal proteins were acquired and reference mass corrected via a dual-spray electrospray ionization (ESI) source. Each scan or data point on the Total Ion Chromatogram (TIC) is an average of 15,000 transients, producing a spectrum every second. Mass spectra were created by averaging the scans across each peak and background subtracted against the first 10 seconds of the TIC. The resulting base peak chromatogram shows very high peak capacity (Figure 6). 119 discrete protein masses were identified; 4 of which were identified as 80S ribosomal proteins. In some cases several different masses of the same protein were identified which correlated with known N- and/or C- terminal processing, highlighted in Table 1.

Figure 6. Base peak chromatogram of ribosomal proteins.

No.	RT [min]	Mass	Protein		No.	RT [min]	Mass	Protein
1	22.4	10943	L37		23	53.9	28680	S6
2	24.2	10943	L37		24	54.5	23191 23647	L14 (native & with mod)
3	28.7	6276	L39		25	55.3	23345	L13a
4	31.7	12321	L36a(L44)		26	55.3	16503	L27a
5	36.7	17279	L26		27	56.2	24015	L15
6	36.7	9399	S27a		28	56.7	9170	S21, N-acetylserine
7	37.9	9270	S27a, cleaved C-term.		29	56.7	14776	L23 N-acetylserine
8	42	18449	L21 NG to KR		30	57.8	15954	S19
9	42.3	17623 17779	L24 & L24 cleaved C-term.		31	60	13284	S20
10	42.3	18448	L21		32	60.8	21527	L18
11	43.8	15667	L27		33	61.4	29464	S4
12	44.6	12122	L36		34	63.2	29862	L7a
13	45.6	15644	L28		35	63.2	16314	S16
14	46.6	8087	L38		36	64.5	29466	S4
15	47.3	12465	L35a		37	64.5	15379	S17
16	47.3 (terminal KK)	23922	L13		38	65.8	17091	S13
17	47.7	23922	L13 (terminal KK)		39	66.8	17629	S18 N-acetylserine
18	49.5	27908	L8		40	67.4	22169	S7 N-acetylmethionine
19	49.5	14164	L34		41	69.5	22460	S9
20	50	18343	S11		42	71.6	14708	S15a
21	51	14421	L35		43	72.5	21893	L9
22	51.6	15465	S24		44	75.3	11772	L22
23	51.6	15466	S24		45	77	30355	L7
24	53.4	15727	L32		46	78.5	26585	S3 N-acetylalanine

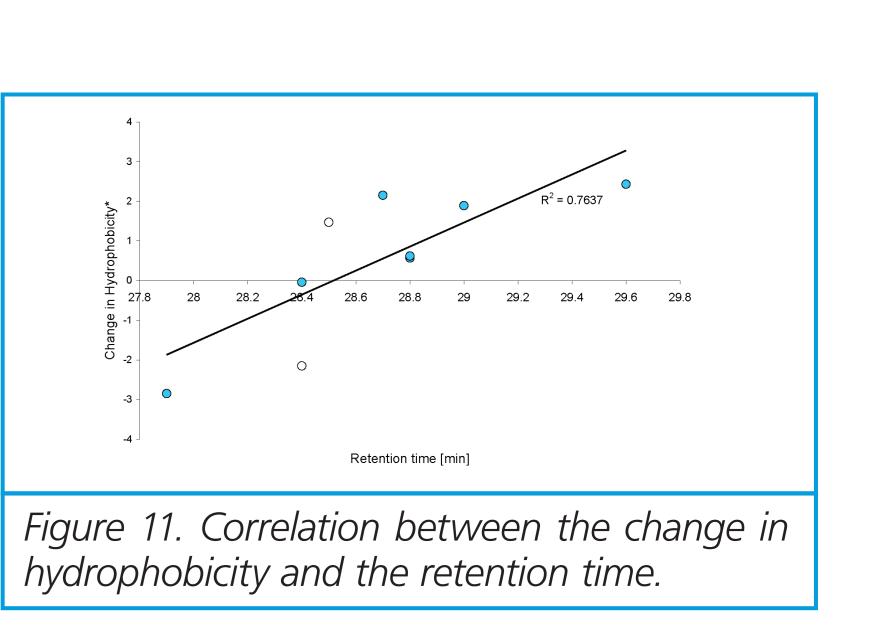Table 1. Chromatogram identification for ribosomal proteins.

Based on the excellent resolution of the intact ribosomal proteins we anticipated the column would be suitable for the analysis of single amino acid mutations in recombinant proteins. The results for the single amino acid mutations intact proteins are shown in Table 2, and Figure 7.

Figure 7. Chromatograms of different mutant proteins.

LC-MS Data for the Mutated Proteins

Mutation	Mw (calc.)	Mw (meas.)	RT [min]	Hydrophobicity
R230Q	33036.8	33037.2	28.8	0.57
W287R	33034.8	33035.9	27.9	-2.85
T73A	33036.1	33036.1	28.8	0.62
S254L	33090.9	33092.5	29.6	2.43
D194G	33006.8	33008.8	28.7	2.15
K107A	33007.7	33009.7	29.0	1.89
Y108A	32972.7	32974.7	28.4	-0.04
S136P	33074.9	33251.2	28.5	1.47
G69D	33122.9	33299.0	28.4	-2.15

Table 2. Properties and retention data for the nine mutant proteins.

The shift in retention time can be correlated to the change in hydrophobicity. A hydrophobicity scale published by Cowan and Whittaker was used because it is based on HPLC retention data[2]. As shown in Figure 11, there is a good correlation between the change in hydrophobicity and the retention time. The open circles indicate mutated proteins where the measured mass differs significantly from the predicted mass.

Figure 11. Correlation between the change in hydrophobicity and the retention time.

Conclusion

The separation of proteins by reversed phase chromatography remains an uncommon technique due to the poor recovery of proteins from the commonly used 300 Å pore size stationary phases. However, with the increasing resolving power of the modern mass spectrometer and the inherent compatibility of the reversed phase solvent systems with MS analysis, we explored the potential of wide pore silicas for the accurate mass analysis of intact ribosomal proteins. We demonstrate excellent resolution of the rat liver 80S ribosome, where in some cases several different masses of the same protein were identified. Further, we were able to separate variations of a 33 kDa protein which differed by a single amino acid mutation.

The change of a single amino acid influences the hydrophobicity and this change correlates well with the shift in retention time of the intact protein.

In the separation of large molecules, especially in an adsorption mode such as reversed phase chromatography, it is important to have unrestricted access to the surfaces inside the pores even when an analyte molecule is already adsorbed at the pore entrance.

References

1. Williamson et al. Eur. J. Biochem, 1997; 246: 786-793

2. R. Cowan and R.G. Whittaker. Peptide Research,1990; 3: 75-80

Acknowledgements

SGE Analytical Science

www.sge.com