Analysis of Intact Proteins Using Liquid Chromatography

R.P. Freeman, H.J. Wirth, A.A. Gooley
SGE Incorporated, 2007 Kramer Lane, Austin Texas. 78758, USA

Abstract

The demand for separation techniques for intact proteins is increasing with the introduction of a new generation of high resolution mass spectrometers which are able to measure the mass of small to medium size proteins very accurately. Liquid chromatography is a valuable tool for separating these proteins prior to the MS analysis. Intact protein chromatography is most commonly used in a top-down approach in proteomics and to determine expression levels during recombinant protein synthesis.

The size of the protein molecule results in very low diffusion coefficients and therefore slow mass transfer in and out of the pore system. A sufficiently large pore diameter is required to minimise the effects of restricted pore diffusion. We show examples of the separation of intact proteins on a column packed with 3 µm C8 silica with 1000 Å pore size. The molecular weight of the protein examples reported here cover ribosomal proteins (≤40 kDa), monoclonal antibodies (~150 kDa) and intact membrane proteins derived from mouse liver.

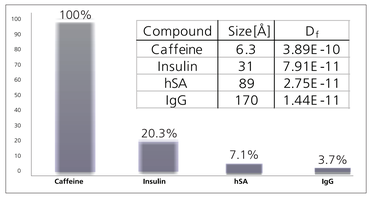
Physico-Chemical Background

Diffusion in Free Solution

The diffusion rate is described by: $D = \frac{k_B T}{6\pi\mu r}$

Einstein-Stokes Equation

D_f = Diffusion rate constant in free solution
k_B = Boltzmann constant
T = Temperature
r = Stokes radius
µ = Viscosity



Compound	Size[Å]	D_f
Caffeine	6.3	3.89E-10
Insulin	31	7.91E-11
hSA	89	2.75E-11
IgG	170	1.44E-11

Figure 1: Normalized diffusion rates of small, medium and large molecules in free solution

Diffusion Inside Pores

$$D_p = D_f\left(1-\frac{r_s}{r_p}\right)^2\left[1-2.104\left(\frac{r_s}{r_p}\right)+2.09\left(\frac{r_s}{r_p}\right)^3-0.95\left(\frac{r_s}{r_p}\right)^5\right]$$

Renkin Equation (E.M. Renkin, J.Gen.Physio., 38 (1954) 225.).

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

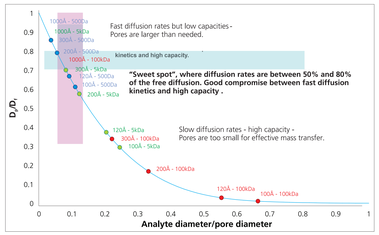


Figure 2: Relative diffusion rates in free solution and inside pores as function of molecule size to pore size ratio

Dedicated Protein Columns - C8, 3 µm, 1000Å

The SGE column range for biological analytes focuses on an optimized pore size for both peptides (200 Å and 300 Å) and proteins (1000 Å) and small particle size (3 µm) to enable lowest possible mass transfer restrictions.
Exposure to the analyte to metal surfaces is minimized through the use of glass-lined or PEEK™-lined stainless steel tubing. All tubing is PEEK™ coated fused silica. Great emphasis is put on the avoidance of void volumes in the design of the capillary column range.

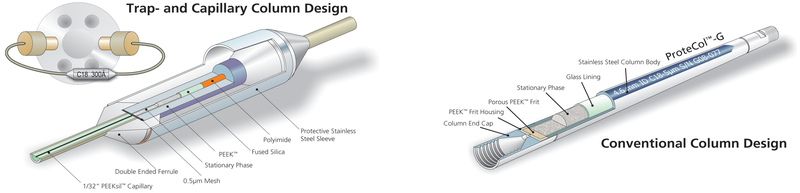


Figure 3: Design of the ProteCol™ Range Columns

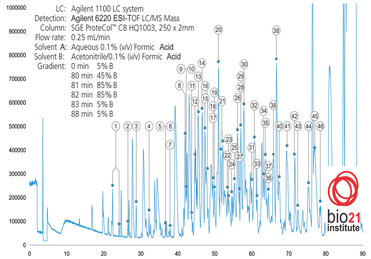
Accurate Mass Analysis of Intact Ribosomal Proteins

Introduction

Since ribosomal proteins are relatively small (Mw=6,000 to 40,000), they can be rapidly identified by accurate LC MS analysis of the intact proteins. In the present application ribosomal proteins isolated from rat liver were separated on a ProteCol™ C8 HQ1003 column.

Sample Preparation

80S ribosomal proteins were isolated from a rat liver microsomal preparation (Williamson et al; 1997, Eur. J. Biochem. 246: 786-793). One optical density unit at 260 nm of 80S ribosomal proteins was mixed with two 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 min at 13,000 rpm and the supernatant was collected into a sample vial ready for LCMS analysis.



No.	RT [min]	Mass	Protein
1	22.4	10943	L37
1	24.2	10943	L37
2	26.5	6648	S30
3	28.7	6276	L39
4	31.9	12321	L36a(L44)
5	36.7	17279	L26
6	36.7	9399	S27a
7	37.9	9270	S27a; cleaved C-term.
8	42	18449	L21 NG to KR
9	42.3	17623 17779	L24 & L24 (cterm)
10	42.3	18448	L21
11	43.8	15667	L27
12	44.6	12122	L36
13	45.6	15644	L28
14	46.6	8087	L38
15	47.3	12465	L35a
16	47.3	23922	L13 (terminal KK)
16	47.7	23922	L13 (terminal KK)
17	49.1	27908	L8
17	49.5	27908	L8
18	49.5	14164	L31
19	50	18343	S11
20	51	14421	L35
21	51.6	15465	S24
21	52.2	15466	S24
22	53.4	15727	L32
23	53.9	28680	S6

No.	RT [min]	Mass	Protein
24	54.5	23191 23647	L14 (native & with mod)
25	55.3	23345	L13a
26	55.3	16503	L27a
27	56.2	24015	L15
28	56.7	9170	S21, N-acetylmethionine
29	56.7	14776	L23 N-acetylserine
30	57.8	15954	S19
31	60	13284	S20
32	60.8	21527	L18
33	61.4	29464	S4
34	63.2	29862	L7a
35	63.2	16314	S16
35	63.6	16314	S16
36	64.5	29466	S4
37	64.5	15379	S17
38	65.8	17091	S13
39	66.8	17629	S18 N-acetylserine
40	66.8	22169	S7 N-acetylmeth
40	67.4	22169	S7 N-acetylmeth
41	69.5	22460	S9
42	71.6	14708	S15a
43	72.5	21893	L9
44	75.3	11772	P2
44	77	11852	P2
45	77	30355	L7
46	78.5	30356	L7
46	78.5	26585	S3 N-acetylalanine

Figure 4: Base peak chromatogram of ribosomal proteins

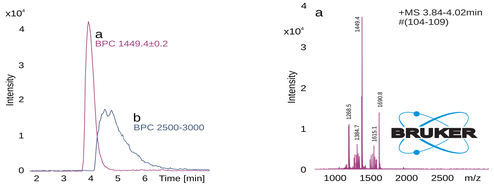
Data Analysis and Results

All data were acquired and reference mass corrected via a dual-spray electrospray ionisation (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 - 119 discrete protein masses were identified; 46 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.

Accurate Mass Determination of Intact Monoclonal Antibodies

A 10 mm x 2 mm ID trap column was used for desalting and sample focussing prior to the MS analysis. The trap column was packed with 3 µm - 1000 Å C8 silica. Again, the large pore size of the stationary phase facilitates a narrow elution profile.

LC System: UltiMate 3000 RSLC-Ti
RP Column: ProteCol™ C8 HQ1008, 10 mm x 2 mm, 3 µm, 1000 Å (trap column)
Mobile phase A: 100% water + 0.1% Formic Acid
Mobile phase B: 5/95 water/acetonitrile, 0.1% Formic Acid
Gradient: 5-90%B in 10 minutes
Flow-rate: 400 µl/min
Temperature: 70°C
Sample amount: mAb 20 mg/mL, 5 µL injected
MS: Bruker maXis



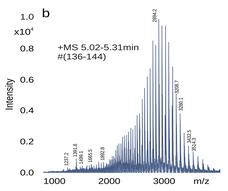


Figure 5: Base peak chromatogram of mAb and MS traces

Analysis of Membrane Proteins

The experiment was performed in three steps:
1. Separation of intact membrane protein sample derived from mouse liver using a 3 µm C8 column with 1000 Å pore size.
2. Fractions collected from 1. were digested with trypsin off-line.
3. Digested fractions were analyzed by LC-MS/MS using 3 µm C18 column with 300 Å pore size.

Chromatographic conditions for analytical separation

System: Agilent 1100 CapLC with Agilent MSD-iontrap MS
Column: ProteCol™-C8 HQ1003 3 µm; 1000 Å 150 mm x 300 µm ID
Sample: 3µl AOHUPO-MPI standard
Flow rate: 5.0 µl/min
Temperature: 80°C
Mobile Phase A: 0.1% formic acid in water
Mobile Phase B: 0.09% formic acid in acetonitrile

1-minute fractions of the intact protein separation were collected and digested with trypsin off-line. The dried fractions were reconstituted in 100 mM NH_4HCO_3 and digested with trypsin overnight, acidified with 1% formic acid (FA), concentrated and re-diluted to 10 µl with 1% FA.

Chromatographic conditions
System: TSP4000 pump, Surveyor autosampler, Thermo linear ion trap
Column: ProteCol™-C18 HQ303 150 µm ID x 100 mm
Sample volume: 10 µl
Mobile phase A: 0.1% FA in 5% acetonitrile
Mobile phase B: 0.1% FA in 90% acetonitrile
MS: 400-1500 mass range; top 6 ions fragmented with 39% collision energy

Data analysis

Raw data files were converted to mzXML and searched against the Ensembl mouse database using Xtandem algorithm (GPM-XE software).
The found proteins were statistically filtered either by number of identified peptides (>4) and/or by a cutoff score (10^{-10}). As a result 542 proteins were identified with high confidence.

Conclusions

By combining stationary phases with the right surface chemistry and the appropriate pore size for the analyte it is possible to separate even "difficult" samples such as membrane proteins. Our investigations into the effect of particle porosity led to the following recommendation for analyte size:

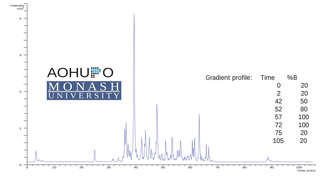


Figure 6: Base peak chromatogram of membrane proteins on a capillary column

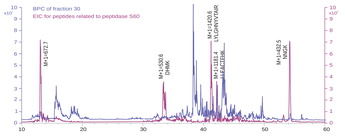
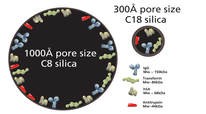


Figure 7: Example of analysis of digested fractions (fraction 30)



Pore Size	Spherical lower	Analyte upper	Cylindrical lower	Analyte upper
100	30	800	13	350
120	50	1400	23	610
200	240	6,400	100	2,800
300	800	21,400	350	9,500
1000	29,400	800,000	13,000	350,000

Figure 8: Relative pore and protein sizes for 1000 Å and 300 Å pores

Acknowledgements

The ribosomal protein results were obtained from Nicholas Williamson and Paul O'Donnell from Bio21 Mass Spectrometry and Peptide Institute, Bio21 Molecular Science and Biotechnology, Melbourne.
Monoclonal antibody data was provided by Matthias Pelzing, Bruker Biosciences, Melbourne.
The membrane protein analysis was done in collaboration with Mibel Aguilar, Monash University and Paul Haynes, APAF, Australia.