

Rapid Identification and Screening of Proteins from Whole Cell Lysates of Human Erythroleukemia Cells in the Liquid Phase, Using Non-porous Reversed Phase High-performance Liquid Chromatography Separations of Proteins Followed by Multi-assisted Laser Desorption/Ionization Mass Spectrometry Analysis and Sequence Database Searching

Yajuan Chen, Dan Wall and David M. Lubman*

Department of Chemistry, The University of Michigan, 930 N. University Ave. Ann Arbor, MI 48109-1055, USA

SPONSOR REFEREE: D. Desidero, Department of Neurology, The University of Tennessee Memphis, Memphis, TN 38163, USA

Non-porous reversed phase (NPRP) high-performance liquid chromatography (HPLC) has been used as a rapid method to separate proteins from whole cell lysates of human erythroleukemia (HEL) cells. Using phosphate-buffered saline (PBS) as a lysis buffer to extract proteins from HEL cells, more than 100 proteins of molecular weight up to 30 kDa were separated by the NPRP HPLC method, using a programmed acetonitrile:H₂O gradient. The separated proteins were collected as liquid fractions as they eluted, and were further separated on the NPRP column with a different gradient to separate coeluting peaks. The isolated protein fractions were analyzed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) to determine the molecular weight of the protein. The proteins were cleaved by chemical or enzymatic digestion to produce peptide maps, which were analyzed by pulsed delayed extraction MALDI-MS. The peptide maps were matched against a database search to determine the protein identity. In some cases, several enzymes were used in order to find exactly one match against the database. This methodology is demonstrated for several proteins isolated from HEL cells and identified via database matching.

© 1998 John Wiley & Sons, Ltd.

Received 2 October 1998; Revised 8 October 1998; Accepted 12 October 1998

An important goal in the field of proteomics involves the ability to rapidly analyze and identify the entire measurable protein profile in cells, tissues or even small organisms.¹ Moreover, methodology will be needed that can profile the total protein composition of a cell and detect even minor alterations or modifications in protein structure that might result in changes in the function of that protein in the cell. Such protein profiling methods should provide protein analysis where the identification of target proteins and their modified versions can be rapidly achieved. Further, such profiling must be able to provide information on many proteins so that changes in cell mechanisms that involve changes in several proteins in terms of expression and structure can be monitored simultaneously.

*Correspondence to: D. M. Lubman, Department of Chemistry, The University of Michigan, 930 N. University Ave. Ann Arbor, MI 48109-1055, USA.

Contract/grant sponsor: National Science Foundation; Contract/grant number: DEB 912006.

Contract/grant sponsor: US Army ERDEC; Contract/grant number: DAAD05-98-P-0796.

Contract/grant sponsor: National Institutes of Health; Contract/grant number: R01GM49500.

The primary method used to study changes in cellular protein expression has been 2-D polyacrylamide gel electrophoresis (2-D PAGE).^{2–6} This method can separate over a thousand protein spots, and the spot pattern is characteristic of each cell type. The 2-D gel spot pattern provides a reference in which a change in the pattern is indicative of abnormal changes in a cell or tissue due to disease or mutation. The pattern is also indicative of normal changes such as aging, stress, etc. The 2-D gel method, however, provides a separation of the cell components and an approximate molecular weight; exact protein identification must still be provided by either Edman sequencing or molecular weight measurements provided by mass spectrometry. In recent research, various membrane-blotting methods^{7–10} and in-gel digestion methods^{11–21} have been combined with mass spectral analysis to provide peptide maps of gel protein spots for identification. Nevertheless, alternative methods of cellular protein analysis are desirable because the 2-D gel method is labor intensive, slow and not readily automated. Moreover, the loadability of sample onto the gel is limited, which means that proteins expressed at low levels may be difficult to detect and analyze.

As an alternative to 2-D gel separation, we have

developed a non-porous reversed phase (NPRP) HPLC method for rapid separation of proteins from whole cell lysates.^{22–24} The use of NPRP HPLC allows a means of separating and isolating proteins in the liquid phase. This method provides high efficiency recovery of proteins available in cells in low amounts such that further analysis is possible. The method circumvents an inherent problem of analyzing proteins that become embedded in polyacrylamide gels. In addition, the relatively high loadability of RP-HPLC methods should, in principle, provide a substantial amount of material for further analysis by mass spectrometry methods. Also, the use of liquid separations and fraction collection should allow future development of these methods towards automation.

The use of NPRP HPLC for separation of whole cell lysates is a key issue in this work. In previous research, a traditional porous RP C-18 column was used to separate proteins from whole lysates of bacterial cells^{25,26} and from human lymphocyte nuclei²⁷ for detection and fingerprinting by MALDI-MS. The use of non-porous, C-18 coated, silica-based packing materials, however, has been shown to provide distinct advantages in the separation of proteins compared to conventional porous materials. The separation of complex protein mixtures in cells can be accomplished in one third of the time required for porous packed column separations with much enhanced protein recovery and resolution. The use of nonporous packing prevents proteins from sticking to the inside of the pores of porous materials, thus allowing a much greater speed of the separation, a decrease in chemical noise background and enhanced protein recovery.²⁴

In this research, we demonstrate the rapid separation of proteins from whole cell lysates of human erythroleukemia (HEL) cells, using NPRP HPLC. It is demonstrated that, using a phosphate-buffered saline (PBS) as a lysis buffer to extract proteins from HEL cells, more than 100 proteins of molecular weight up to 30 kDa could be separated by gradient NPRP HPLC and collected in the liquid phase. In addition, the rapid separation and efficient recovery of proteins using the NP phase allows for a rapid second separation by the NP column with a modified gradient to separate coeluting peaks so that isolated protein fractions are obtained. It will be shown that the isolated proteins can be digested by CNBr or trypsin to produce peptide maps that can be analyzed by pulsed delayed extraction (PDE) MALDI-MS.^{28–33} The peptide maps can be matched against a database search to determine the protein identity. This methodology is demonstrated for several proteins isolated from HEL cells and identified via database matching.

EXPERIMENTAL SECTION

Cell cultures

The human erythroleukemia (HEL) cell line was obtained from the Department of Pediatrics and Surgery, the University of Michigan. HEL cells were cultured (7% CO₂, 37°C) in RPMI-1640 medium (Gibco) containing 4 mM glutamine, 2 mM pyruvate, 10% fetal bovine serum (Gibco) and 250 mg hygromycin (Sigma).

Cell lysates

Approximately 5×10^7 cells were centrifuged at 700 rpm for 5 min at room temperature, and the supernatant

discarded. The cell pellet was subsequently washed three times with cold PBS (Gibco). This process involved vortexing the sample for a few seconds, centrifuging the samples and discarding the supernatant. After the final removal of supernatant, the cells, in a small volume of residual PBS, were sonicated in an ice bath for 10 to 20 seconds. PBS (1 mL) was the buffer for the lysed cells. The mixture was centrifuged at maximum speed, and the supernatant collected. The supernatant was centrifuged again to eliminate any remaining particles. The cell lysates collected were either directly used for analysis or stored at -80°C .

NPRP HPLC of HEL cell proteins

A Beckman (Fullerton, CA, USA) System Gold HPLC was used for this research. It has a programmable solvent delivery module with a dual-pump (Model 127). This module includes a System Control Center display and keypad, which permits direct control of the pump and external modules. The detector was a programmable detector module (Model 166) with a deuterium lamp. The column was purchased from MICRA (Northbrook, IL, USA). It contained non-porous C18 packing material with a 1.5 μm particle diameter. The dimension of the column was 4.6×33 mm, the widebore column was used to collect an adequate amount of material, and the short column allowed for rapid separations in minutes.

The buffers are buffer A: 0.1% TFA (Sigma), H₂O (MQ), and buffer B: 0.1% TFA, acetonitrile (HPLC grade, Sigma-Aldrich). About 70 μL of the HEL cell lysate was injected into the injection valve, containing a 100 μL sample loop, and introduced into the HPLC system. The separation of proteins from the HEL cell lysate used the following conditions: buffer B ranged from 0% to 20% for 2 min, 20% to 30% for 5 min, 30% to 40% for 10 min, 40% to 50% for 5 min, 50% to 80% for 4 min, 80% to 100% for 2 min, and 100% to 0% for 1 min. The flow rate for all HPLC experiments was 1 mL/min. The column temperature was controlled at 60°C by a column heater (Timberline), and the detection wavelength was 214 nm. Selected peaks from the separation were collected, and were further separated on the NPRP column by adjusting the gradient of buffer B in order to achieve improved separation. The second step separated proteins that could not be separated in the first step.

MALDI-TOFMS analysis

The TOFMS employed in these studies was a modified Wiley–McLaren design with a field-free drift length of 1 m.³⁴ The TOF was capable of high voltage acceleration up to ± 20 kV (R. M. Jordan Co., Grass Valley, CA, USA). The laser source used to produce MALDI was a DCR 11 Nd:YAG laser system (Spectraphysics, San Jose, CA, USA). The laser beam was focused onto the probe tip at a 45° angle to the probe surface with a single 12.5 in. f.l. quartz lens to a spot size of ~ 0.2 mm \times 0.5 mm. The resulting power density at this spot was estimated to be $\sim 5 \times 10^6$ to 1×10^7 W/cm² at 355 nm radiation. The detector was an 18 mm triple microchannel plate (MCP) detector (R. M. Jordan Co.), which adapted a CuBe conversion dynode with post-acceleration (PA) capability up to ± 12 kV in front of the MCP. The total ion acceleration across the TOF device may thus be > 30 kV. The post-acceleration stage enhances the efficiency for detection of

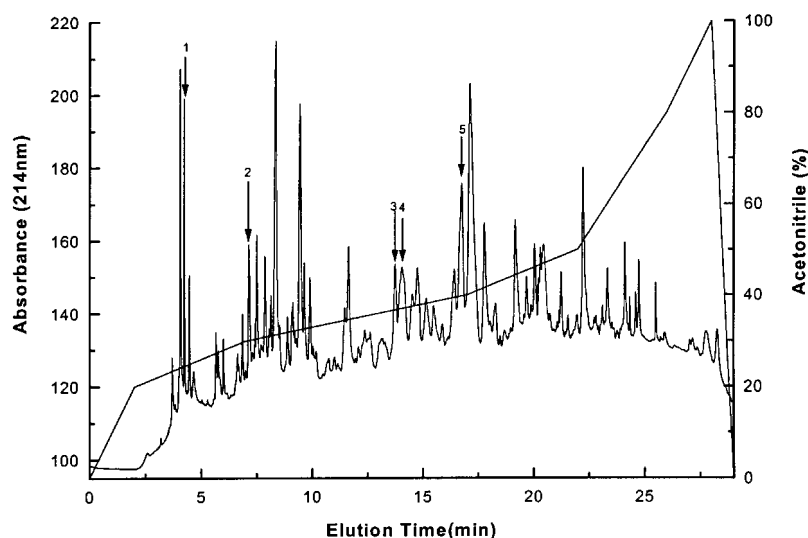


Figure 1. The NPRP HPLC separation of proteins from a HEL cell lysate.

heavy species, but at the expense of resolution. The voltages were similar for all experiments: repeller at +15 kV, extractor plate +12 kV, extractor plate 2 was grounded, and the PA was held at a voltage of -10 to -13 kV. The TOF flight tube was pumped to a base pressure of 2×10^{-7} Torr, with a Turbo-V250 pump (Varian Associates). Data were recorded with a LeCroy 9350 M digital oscilloscope (LeCroy Corp., Chestnut Ridge, NY, USA), and processed on a Gateway 586 166 MHz computer. PDE³⁵⁻³⁸ as described in previous work was used for the trypsin and CNBr digest analysis. PDE used a simple high voltage transistor switch constructed as described.³⁹ The switch provided a 0-3 kV voltage pulse with a 75 ns fall time.

The matrix was a saturated solution of α -cyano-4-hydroxycinnamic acid (α CHCA, Sigma-Aldrich) in acetone (Sigma-Aldrich) +1% TFA. The protein fractions collected from the NPRP HPLC separation were dried on a Speed-Vac. A solution (5 μ L), consisting of 50% acetonitrile +50% H₂O (each containing 0.1% TFA), was added to dissolve the protein samples. Sample solution (1.5 μ L) was loaded onto the stainless steel probe tip, followed by 2 μ L of matrix solution. These solutions were mixed well and allowed to dry prior to MALDI analysis. In order to improve mass accuracy, a standard mixture of substance P (MW 1347.6), insulin (MW, 5733.5) and horse heart cytochrome C (MW 12 384) was spotted next to the sample spot on the probe.

Protein digestion and MS analysis

In order to identify the protein fractions collected from HPLC separations, chemical and enzymatic reagents were used to produce peptide maps. CNBr and trypsin were used to generate these peptide maps. In the case of CNBr digestion, 2 μ L of 0.15 M CNBr in 70% formic acid was added to the protein samples, which contain 8 μ L of 50% acetonitrile and 0.1% TFA. The samples were incubated overnight at 37°C. After digestion was complete, 2 μ L of the solution was applied directly to the probe tip, followed by 2.5 μ L of CHCA matrix solution; the matrix was air dried. In the case of tryptic digestion, 20 μ L of 50 mM ammonium bicarbonate was added to each dried protein sample, followed by 2 μ L of 0.1 μ g/ μ L trypsin in 1 mM HCl.

The sample was incubated for about 20 hours at 37°C. After the digestion was complete, the sample was dried by Speed-Vac to eliminate the ammonium bicarbonate. About 5 μ L of 50% acetonitrile (0.1% TFA) was used to dissolve the protein digest sample. The sample was analyzed by MALDI-MS analysis.

Database search for protein identification

The peptide maps obtained by digestion for each protein were searched directly by using the MS-Fit sequence database search tool in the Protein Prospector program, which was designed at UCSF. The program is available on the Internet (<http://prospector.ucsf.edu>). In order to narrow the search range, protein species, molecular weight range and pI are usually required as input into the database. In our case, the unknown proteins were obtained from HPLC separation. No pI information was available. Because the proteins were from a HEL cell extract, *Homo Sapiens* was chosen as the species. The molecular weight of each protein was determined by MALDI analysis. In the database analysis, a relatively wide MW range of up to 500 Da was chosen so that the protein was not eliminated from the search due to a shift of the MW from its theoretical value by post-translational modifications or other differences in sequence. The tolerance for the search of peptides against the database was set at 50 Da.

In the case of the peptide search, the CNBr digest usually generates a few large peptide fragments. The result is that fewer matches are found during the database search compared with those from a trypsin digest. Some searches found exactly one match for the protein CNBr digestion. For those searches in which more than one match were found, the identity of the proteins could be obtained by finding the common match of both CNBr and tryptic digest database searches. If still more than one match was obtained, then additional proteases could be used to find a single common match.

RESULTS AND DISCUSSION

A separation of a HEL cell lysate by NPRP HPLC is shown

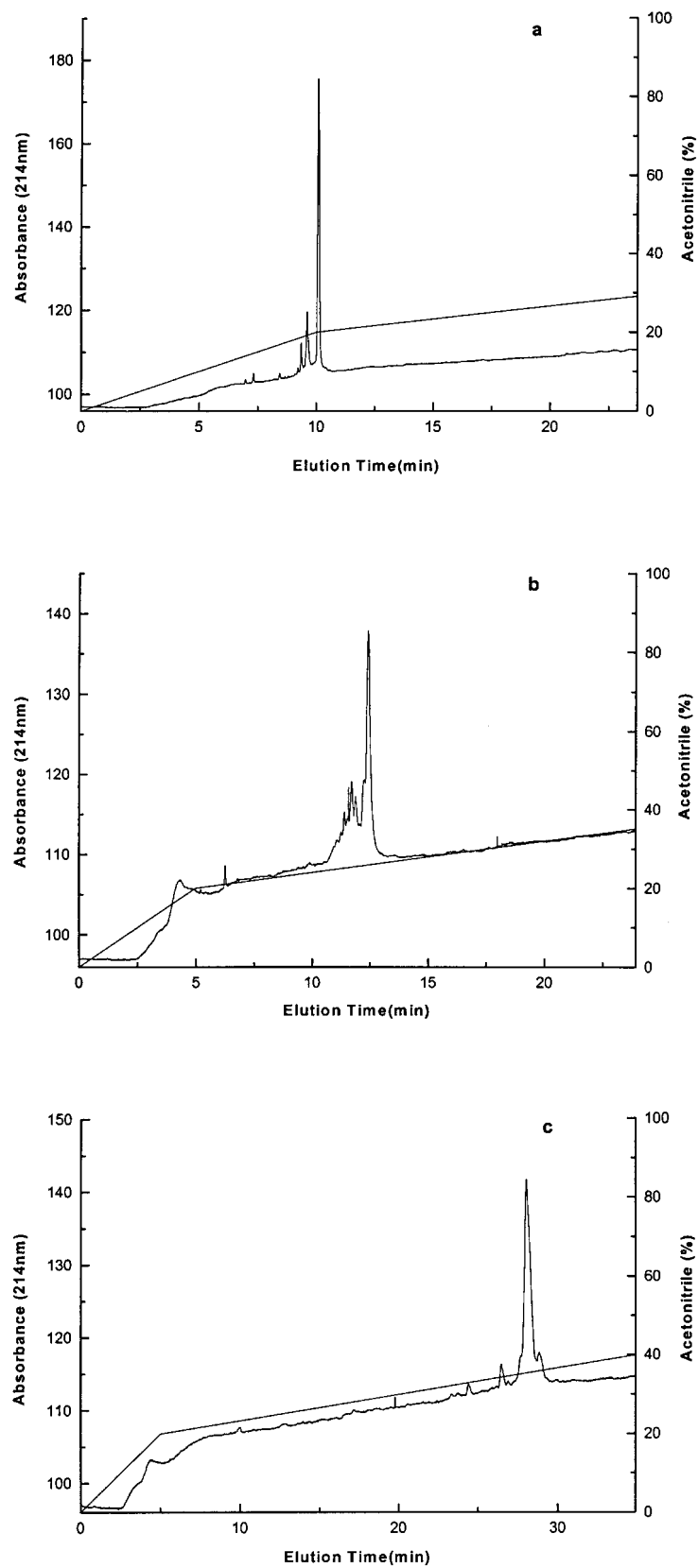


Figure 2. Peaks 1, 2, and 3 were collected as the eluent of the separation in Fig. 1, and separated a second time as shown in (a), (b) and (c), respectively. In each case, the dominant peak of each chromatogram was collected for further analysis. The gradient for buffer B was: (a) 0% to 20% in 10 min, 20% to 30% in 15 min.; (b) 0% to 20% in 5 min, 20% to 40% in 25 min.; (c) 0% to 20% in 5 min, 20% to 40% in 30 min.

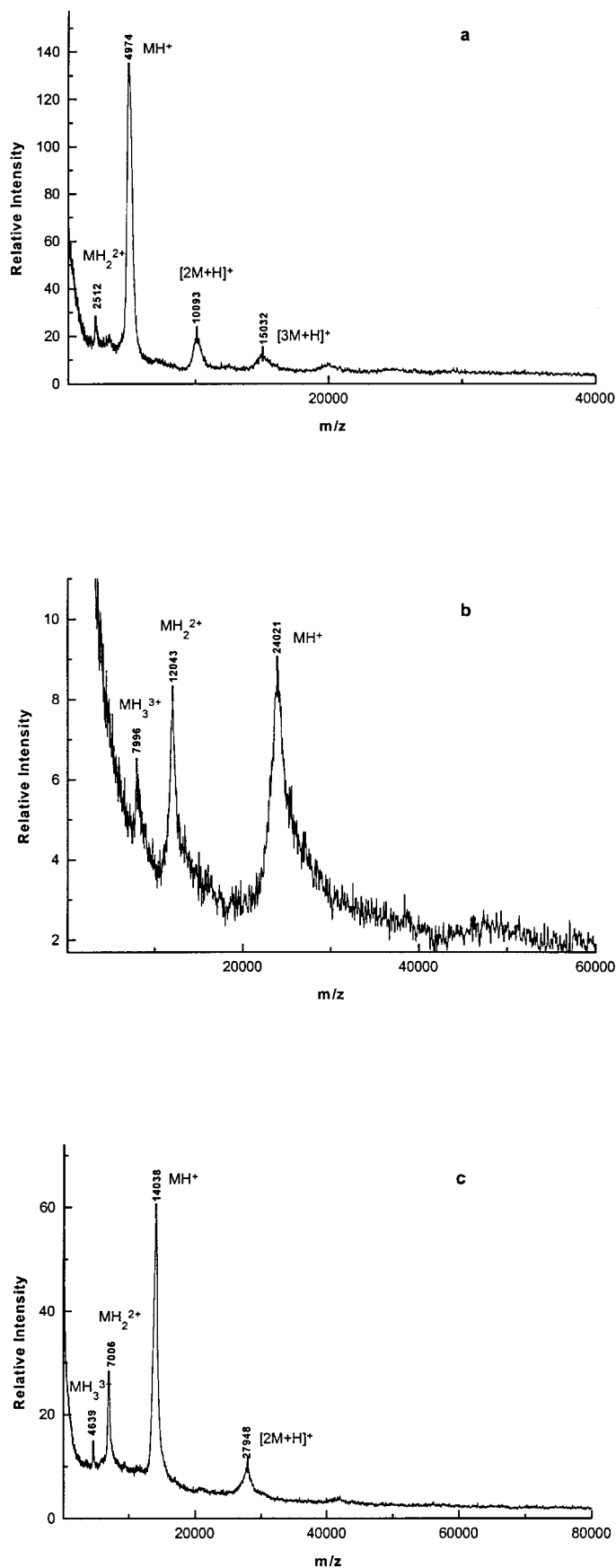


Figure 3. MALDI-MS spectra of the three isolated proteins from Fig 2(a)–(c) as shown in (a)–(c), respectively. The spectra were obtained from the average of 50 laser pulses.

Table 1. Measured molecular weights for the five HPLC-separated, HEL proteins, analyzed by MALDI mass spectrometry

Peak No.	Protein name	Measured MW(Da)	Theoretical MW(Da)
1	thymosin beta 4 Y isoform	4974	5013
2	high mobility group protein 2	24 021	24 034
3	putative DNA binding protein	14 038	14 046
4	ribonucleoprotein	11 431	11 454
5	calmodulin	17 207	17 163

in Fig. 1. In order to achieve a rapid, efficient separation of the cell extract, a short column, 4.6×33 mm, packed with $1.5 \mu\text{m}$ nonporous C18 was used. The nonporous columns have the advantage of eliminating nonspecific adsorption and have been found to be most suitable for the rapid separation of large biomolecules, such as proteins. A column heater provided an elevated temperature of 60°C in order to enhance the speed and quality of the separation. Under the conditions used in this work, the separation of

proteins of up to 30 kDa from the HEL cell lysate was achieved in less than 30 minutes.

In this research a PBS extraction buffer provided improved results as compared to strong reagents for extracting proteins from the cells such as 6 M urea or 8 M guanidium HCl, which are often recommended for this purpose. Using the PBS extraction method and the gradient NPRP HPLC method outlined herein, over 100 UV detected proteins could be separated by the HPLC method. Proteins of up to ~ 30 kDa were separated and detected using this procedure. Each peak shown in Fig. 1 could be collected and analyzed by MALDI-MS. Many of the peaks in Fig. 1 may contain more than one major protein, as determined by MALDI-MS. In order to isolate individual proteins, a second separation of selected peaks in Fig. 1 could be performed, using a lower percentage of organic modifier per minute gradient to separate the protein components. A second separation of several peaks from Fig. 1 with collection of the major protein in each fraction is shown in Fig. 2.

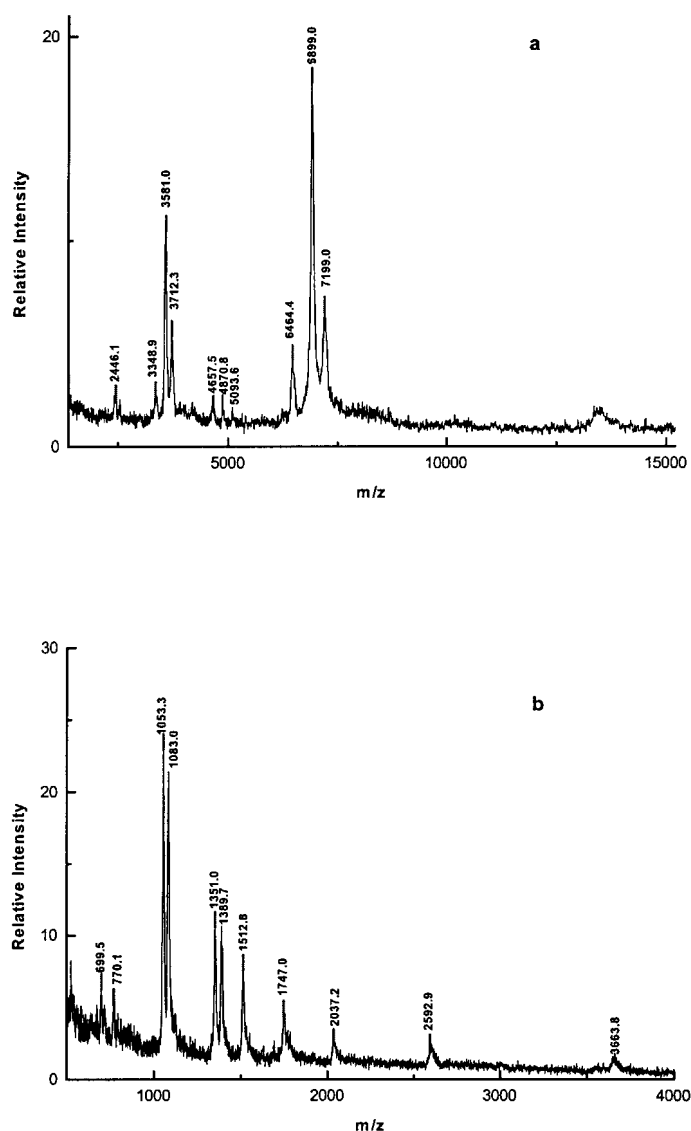


Figure 4. The peptide mappings of two isolated proteins: (a) CNBr digest of the protein isolated from Fig. 2; (b) tryptic digest of a second HPLC separation of a peak from Fig. 1. The spectra were obtained from the average of 50 laser pulses.

Table 2. Database search results for CNBr and tryptic digests of the five proteins separated from HEL cell extract by HPLC, identified by MALDI mass spectrometry

Peak No.	M _{meas} (u)	M _{calc} (u)	ΔM(u)	residues	amino acid sequence
1	(CNBr digest)				
	573.0	586.2	-13.2	2-7	(M)SDKPGh(A)
	4271.2	4264.2	7.0	8-44	(M)AEIEKFDKSKLKKTTETQ EKNPLSSKETIEQERQAGES(-)
	(Tryptic digest)				
	654.7	645.4	9.3	27-32	(K)NPLSSK(E)
	876.3	865.5	10.7	13-19	(K)FDKSKLK(K)
	1079.0	1104.6	-25.6	18-26	(K)LKKTETQEK(N)
	1308.2	1319.7	-11.5	16-26	(K)SKLKKTTETQEK(N)
	1345.8	1335.6	10.2	1-12	(-)MSDKPGMAEIEK(F)
	1699.0	1709.9	-10.9	13-26	(-)MSDKPGMAEIEKFDK(S)
	1874.5	1867.9	6.6	1-17	(-)SDKPGMAEIEKFDKSK(L)
	2303.3	2310.2	-6.9	1-20	(-)MSDKPGMAEIEKFD KSKLKK(T)
	2816.4	2831.5	-15.1	16-39	(K)SKLKKTTETQEKNPLSS KETIEQER(Q)
	2	(CNBr digest)			
5944.4		5958.9	-14.5	1-52	(-)GKGDPNPKPRGKM*SSYAF FVQTCREEHKKKHPDSSVN FAEFSKKCSERWkTh(S)
6303.9		6281.4	22.5	76-132	(M)KNYVPPKGDKKGKKKD PNAPKRPPSAFFLFCSEHRPKI KSEHPGLSIGDTAKKLGEh(W)
7738.8		7732.1	6.7	64-132	(M)AKSDKARYDREMKNY VPPKGDKKGKKKDPNAP KRPPSAFFLFCSEHRPKIKS EHPGLSIGDTAKKLGEh(W)
8891.9		8917.1	-25.2	133-209	(M)WSEQSAKDKQPYEQKA AKLKEYEKDIAAYRAGK SEAGKKGPRPTGSKKKNEP EDEEEEEDEDEDEDEE DEE(-)
15077.0		15073.6	3.4	1-132	(-)MGKGDPNPKPRGKMSSY AFFVQTCREEHKKKHPDSS VNFAEFSKKCSERWkTMSA KEKSKFEDMAKSDKARYD REMKNYVPPKGDKKGKKK DPNAPKRPPSAFFLFCSEHR PKIKSEHPGLSIGDTAKKL GEh(W)
(Tryptic digest)					
717.0		717.4	-0.4	77-82	(K)NYVPPK(G)
864.8		867.4	-2.6	49-55	(R)WKTMSAK(E)
992.5		984.6	7.9	174-183	(K)GPRPTGSKK(K)
1172.5		1195.6	-23.1	1-12	(-)GKGDPNPKPRGK(M)
1331.0		1327.7	3.3	155-165	(K)YEKIDIAAYRAK(G)
1481.2		1455.6	25.6	13-24	(K)M*SSYAFFVQTCR(E)
1602.9		1599.8	3.1	45-57	(K)CSERWkTM*SAKEK(S)
1731.4	1731.8	-0.4	60-73	(K)FEDMAKSDKARYDR(E)	
1943.7	1943.0	0.7	44-59	(K)KCSERWkTMSAKE KSK(F)	
2027.0	2011.1	15.9	151-167	(K)LKEYEKDIAAYR AKGK(S)	
2850.7	2847.4	3.3	25-48	(R)EEHKKKHPDSSVNFA EFSKKCSER(W)	
3384.9	3384.2	0.7	183-209	(K)KKNPEDEEEEEEE DEDEEEEEDEE(-)	
3	(CNBr digest)				
	2446.1	2451.9	-5.8	1-20	(-)MMSSYERSMMSYERS MMSPh(A)*
		2456.0	-9.9	21-41	(M)AERSMMSAYERSMMS AYERSh(M)
	3348.9	3348.4	0.5	42-70	(M)MSPMAERSMMSAYE RSMMSAYERSMMSPh(A)
	3581.0	3593.5	-12.5	46-76	(M)AERSMMSAYERSMMS AYERSMMSPMADRSMh(S)

Table 2. continued

Peak No.	M _{meas} (u)	M _{calc} (u)	ΔM(u)	residues	amino acid sequence	
3	3712.3	3707.5	4.8	86–119	(M)SSYSAADRSMSSY SAADRSMSSYTAD RSMSh(A)*	
	4657.5	4657.9	-0.4	77–119	(M)SMGADRSMSSYSA ADRSMSSYSAADR MMSSYTADRSMSh(A)	
	4870.8	4875.0	-4.2	43–85	(M)SPMAERSMMSAYERS MMSAYERSMSPMAD RSMMSMGADRSMh(S)	
	5093.6	5089.1	4.5	3–45	(M)SSYERSMMSYERSM MSPMAERSMMSAYERSM MSAYERSMMSPh(A)*	
	6464.4	6465.6	-1.2	60–117	(M)SAYERSMMSPMADR SMMSMGADRSMSS YSAADRSMSSYSA ADRSMSSYTADR Mh(S)*	
	6899.0	6896.8	2.2	1–58	(-)MMSSYERSMMSYERSM MSPMAERSMMSAYER SMMSAYERSMSPM AERSMMSAYERSh(M)*	
	7199.0	7188.9	10.1	17–78	(M)MSPMAERSMMSAY ERSMMSAYERSMSP MAERSMMSAYERSMM SAYERSMSPMADR MMSH(G)	
	4	(CNBr digest)				
		792.7	808.4	-15.7	89–97	(M)MPVGPAPGh(R)
1977.7		1979.9	-2.2	78–97	(M)MPMMGPPPGMMP VGPAPGh(R)*	
2806.8		2804.3	2.5	79–106	(M)PMMGPPPGMMPV GPAPGMRPPMGICQ(-)*	
3032.8		3031.3	1.5	78–106	(M)MPMMGPPPGMMPV GPAPGMRPPMGICQ(-)*	
3728.6		3730.8	-2.2	53–88	(M)IPPPSLPGPPRPGM MPTPIWGPSMMPMMG PPPGh(M)*	
4606.5		4600.2	6.3	53–97	(M)IPPPSLPGPPRPGMMP TPIWGPSMMPMMGPPP PGMMPVGPAPGh(R)*	
5611.1		5603.6	7.5	53–106	(M)IPPPSLPGPPRPGMM PTPIWGPSMMPMMGPP PPGMPVGPAPGMRPP MGGICQ(-)*	
9130.8		9125.3	5.5	19–106	(M)EEQAQSLIDKTTAAF QQGKIPPTPFSAPPAGA MIPPPSLPGPPRPGMMP TPIWGPSMMPMMGPPP GMMPVGPAPGMRPPMG GICQ(-)*	
11141.5		11116.4	25.1	1–101	(-)RSRRKHKENVKDYQ KWMEEQAQSLIDKTTAA FQQGKIPPTPFSAPPAG AMIPPPSLPGPPRPGMMP TPIWGPSMMPMMGPPP GMMPVGPAPGMRPPh(G)*	
5		(CNBr digest)				
	2827.1	2791.3	35.8	129–152	(M)IREADIDGDGVNYE EFVQMMTAK(-)*	
	4049.9	4061.0	-11.1	1–37	(-)MADQLTEEQIAEFKEA FSLFDKDGDTITTKE LGTVh(R)	
	5775.4	5762.7	12.7	1–52	(-)MADQLTEEQIAEFKE AFSLFDKDGDTITTKE LGTVMRSLGQNPTEAEL QDh(I)*	
8005.1	7971.6	33.5	81–149	(M)KDTDSEEEIREAFRVF DKDNGYISAAELRHVMT		

Table 2. continued

Peak No.	M _{meas} (u)	M _{calc} (u)	ΔM(u)	residues	amino acid sequence
5					NLGEKLTDEEVDEMIREA DIDGDGQVNYEEFVQMh(T)* (M)ARKMKDITDSEEEIR EAFRVFDKDGNGYISAAE LRHVMTNLGEKLTDEEV DEMIREADIDGDGQVNY EEFVQMh(T)*
	8479.2	8473.9	5.3	77–149	
	(Tryptic digest)				
	802.8	805.4	–2.6	32–38	(K)ELGTVMR(S)
	1236.8	1265.6	–28.8	99–110	(K)DGNGYISAAELR(H)
	1579.3	1596.7	–17.4	82–94	(K)DITDSEEEIREAFRVFD
	1759.6	1754.9	4.7	95–110	(R)VFDKDGNGYISAA ELR(H)
	2402.8	2391.1	11.7	111–130	(R)HVMTNLGEKLTDE EVDEMIR(E)*
	2795.6	2780.4	15.2	95–119	(R)VFDKDGNGYISAAE LRHVMTNLGEK(L)*
	3335.0	3332.6	2.4	82–110	(K)DITDSEEEIREAFRVFD KDGNGYISAAELR(H)
	3795.6	3820.7	–25.1	120–152	(K)LTDEEVDEMIREAD IDGDGQVNYEEFVQM MTAK(–)
	4089.0	4095.1	–5.9	95–130	(R)VFDKDGNGYISAAEL RHVMTNLGEKLTDEEV DEMIR(E)
	5548.3	5568.6	–20.3	32–81	(K)ELGTVMRSLGQNPT ELQDMINEVDADDLPGNG TIDFPEFLTMMARKMK(D)

* Represents the existence of oxidized methionine in the sequence.

The isolated protein fractions collected were analyzed by MALDI-MS. The MALDI mass spectra of the three isolated proteins collected in Fig. 2(a)–(c) are shown in Fig. 3 (a)–(c), respectively. The three proteins shown range in size from ~5 kDa to 27.5 kDa. Several peaks are observed in each spectrum due to the presence of multiply charged species and also protein dimers, which often result from the use of the α -CHCA matrix. Nevertheless, it appears that only one protein is observed in each fraction and that its molecular weight can be readily determined by MALDI-MS.

In order to identify selected proteins, peptide maps were generated by chemical or enzymatic degradation. The protein digests were analyzed by PDE MALDI-MS (Fig. 4(a) and (b) for a CNBr and tryptic digest of two different proteins. A database search for the results of CNBr and trypsin digestion was performed, using the NCBI database, for the five proteins separated from the cell lysates, using NPRP HPLC as shown in Fig. 1. Each protein fraction was digested by CNBr and trypsin in order to find only one match from the database that corresponded to that protein. The protein identification determined for each protein is shown in Table 1, which lists the observed and theoretical MW's. A summary of the digestion data obtained for each protein with the measured and observed molecular weights and the corresponding amino acid sequence from the database is shown in Table 2. It should be noted that the database program not only provides the intact protein sequence, but also the peptide sequence with modifications such as an oxidized methionine. In Table 2, some of the peptide matches were determined, using these modified proteins, which often provided a better match than the unmodified counterparts.

CONCLUSION

In conclusion, NPRP HPLC has been demonstrated as a means of rapidly separating proteins from whole cell lysates of mammalian cells for collection and analysis of proteins collected in the liquid phase. In this research, HEL cell lines were lysed by a PBS buffer, and over a 100 proteins could be separated by a single NPRP HPLC separation. Proteins up to 30 kDa could be detected. The protein fractions collected could be separated further with a modified gradient in the second separation, using the same C18 NPRP column. The isolated proteins collected from the second separation could be analyzed by MALDI-MS in order to determine the molecular weight. The proteins were digested by CNBr or trypsin to produce peptide maps that could be analyzed by PDE MALDI-MS, and the map compared to the NCBI database to identify the protein. This method has been shown to be a rapid method of separating and identifying proteins in whole cell lysates in the liquid phase without the use of gel electrophoresis separations.

Acknowledgement

We thank Micra Scientific, Inc. for the generous donation of a NPS HPLC column. We also thank Dr. Samir Hanash from the Department of Pediatrics and Surgery, University of Michigan for providing us with the HEL cell line, and Dr. Christina Chang and Dr. David Misek for their technical advice on the handling of these samples.

Credit

We gratefully acknowledge partial support of this work by the National Science Foundation under Grant # DEB 912006 to the Center of Microbial Ecology at Michigan State University, by the U.S. Army

ERDEC under contract #DAAD05-98-P-0796, and by the National Institutes of Health under Grant R01GM49500.

REFERENCES

1. P. Kahn, *Science* **270**, 369–371 (1995).
2. P. H. O'Farrell, *J. Biol. Chem.* **250**, 4007–4021 (1975).
3. S. D. Patterson and R. Aebersold, *Electrophoresis* **16**, 1791–1814 (1995).
4. F. C. Neidhardt, D. B. Appleby, P. Sankar, M. E. Hutton and T. A. Phillips, *Electrophoresis* **10**, 116–122 (1989).
5. N. L. Anderson, R. Esquer-Blasco, J. P. Hofmann and N. G. Anderson, *Electrophoresis* **12**, 907–930 (1991).
6. 'Two-Dimensional Electrophoresis' J. R. Strahler, R. Kuick and S. M. Hanash, in *Protein Structure. A Practical Approach*. T. Creighton (Ed.), IRL Press (1989), pp. 231–266.
7. J. M. Gershoni and G. E. Palade, *Anal. Chem.* **131**, 1–15 (1983).
8. R. Aebersold, J. Leavitt, R. A. Saavedra, L. E. Hood and S. B. Kent, *Proc. Natl. Acad. Sci. USA* **84**, 6970–1974 (1987).
9. J. W. Henzel, T. M. Billeci, J. T. Stults, S. C. Wong, C. Grimely and C. Watanabe, *Proc. Natl. Acad. Sci. USA* **90**, 5011–5015 (1993).
10. X. Liang, J. Bai, Y. H. Liu and D. M. Lubman, *Anal. Chem.* **68**, 1012–1018 (1996).
11. A. Shevchenko, M. Wilm, O. Vorm and M. Mann, *Anal. Chem.* **68**, 850–858 (1996).
12. A. Shevchenko, O. N. Jensen, A. V. Podtelejntkov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie and M. Mann, *Proc. Natl. Acad. Sci. USA* **93**, 14440–14445 (1996).
13. P. L. Courchesne, R. Luethy and S. D. Patterson, *Electrophoresis* **18**, 369–381 (1997).
14. 'Internal Protein Sequencing of SDS-PAGE Separated Proteins: Optimization of In-Gel Digest Protocol' K. Williams, M. L. Presti and K. Stone, In *Techniques in Protein Chemistry VIII*, D. Marshak (Ed.), Academic Press, in press, (1997).
15. G. Li, M. Waltham, N. L. Anderson, E. Unsworth, A. Treston and J. N. Weinstein, *Electrophoresis* **18**, 391–402 (1997).
16. K. L. O'Connell and J. T. Stults, *Electrophoresis* **18**, 349–359 (1997).
17. C. Eckerskorn and R. Grimm, *Electrophoresis* **17**, 899–906 (1996).
18. M. Schuhmacher, M. O. Glocker, M. Wunderlin and M. Przybylski, *Electrophoresis* **17**, 848–854 (1996).
19. N. H. Packer, A. Pawlak, W. C. Kett, A. A. Gooley, J. W. Redmond and K. L. Williams, *Electrophoresis* **18**, 452–460 (1997).
20. R. R. Loo, T. I. Stevenson, C. Mitchell, J. A. Loo and P. C. Andrews, *Anal. Chem.* **68**, 1910–1917 (1996).
21. S. L. Cohen and B. T. Chait, *Anal. Biochem.* **247**, 257–267 (1997).
22. N. Nimura, H. Itoh and T. Kinoshita, *J. Chromatography* **585**, 207–211 (1991).
23. H. Itoh, N. Nimura, T. Kiroshita, N. Nagae and M. Nomura, *Anal. Biochem.* **199**, 7–10 (1991).
24. J. F. Banks and E. E. Gulcicek, *Anal. Chem.* **69**, 3973–3978 (1997).
25. X. Liang, K. Zheng, M. G. Qian and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **10**, 1219–1226 (1996).
26. G. J. Opitek, K. C. Lewis and J. W. Jorgenson, *Anal. Chem.* **69**, 1518–1524 (1997).
27. P. R. Griffin, M. J. MacCoss, J. K. Eng, R. A. Blevins, J. S. Aaronson and J. R. Yates, *Rapid Commun. Mass Spectrom.* **9**, 1546–1551 (1995).
28. C. L. Nilsson, C. M. Murphy and R. Ekman, *Rapid Commun. Mass Spectrom.* **11**, 610–612 (1997).
29. J. R. Yates, A. L. McCormack and J. Eng, *Anal. Chem.* **68**(17), 534A–540A (1996).
30. O. N. Jensen, A. Podtelejntkov and M. Mann, *Rapid Commun. Mass Spectrom.* **10**, 1371–1378 (1996).
31. S. D. Patterson, *Anal. Biochem.* **221**, 1–15 (1994).
32. J. R. Yates, S. Speicher, P. R. Griffin and T. Hunkapiller, *Anal. Biochem.* **214**, 397–408 (1993).
33. J. R. Yates, J. K. Eng, A. L. McCormack and D. Schieltz, *Anal. Chem.* **67**, 1426–1436 (1995).
34. W. C. Wiley and I. H. McLaren, *Rev. Sci. Instrum.* **26**, 1150 (1955).
35. S. M. Colby, T. B. King and J. P. Reilly, *Rapid Commun. Mass Spectrom.* **8**, 865–868 (1994).
36. R. M. Whittal and L. Li, *Anal. Chem.* **67**, 1950–1954 (1995).
37. R. S. Brown and J. J. Lennon, *Anal. Chem.* **67**, 1998–2003 (1995).
38. D. H. Patterson, G. E. Tarr, F. E. Regnier and S. A. Martin, *Rapid Commun. Mass Spectrom.* **9**, 1044–1050 (1995).
39. Y. Zhu, L. He, J. R. Srinivasan and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **11**, 987–992 (1997).