Comprehensive Two-Dimensional High-Performance Liquid Chromatography for the Isolation of Overexpressed Proteins and Proteome Mapping

Gregory J. Opiteck,* Suzanne M. Ramirez,* James W. Jorgenson,*† and M. Arthur Moseley III†

*Department of Chemistry, University of North Carolina at Chapel Hill, Venable Hall, CB 3290, Chapel Hill, North Carolina 27599-3290; and †Department of Analytical Chemistry, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709

Received August 21, 1997

A two-dimensional liquid chromatographic system is described here which uses size-exclusion liquid chromatography (SEC) followed by reversed-phase liquid chromatography (RPLC) to separate the mixture of proteins resulting from the lysis of Escherichia coli cells and to isolate the proteins that they produce. The size-exclusion chromatography can be conducted under either denaturing or non-denaturing conditions. Peaks eluting from the first dimension are automatically subjected to reversed-phase chromatography to separate similarly sized proteins on the basis of their various hydrophobicities. The RPLC also serves to de-salt the analytes so that they can be detected in the deep ultraviolet region at 215 nm regardless of the SEC mobile phase used. The two-dimensional (2D) chromatograms produced in this manner then strongly resemble the format of stained 2D gels, in that spots are displayed on a X-Y axis and intensity represents quantity of analyte. Following chromatographic separation, the analytes are deposited into six 96-well (576 total) polypropylene microtiter plates via a fraction collector. Interesting fractions are analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF/MS) or electrospray mass spectrometry (ESI/MS) depending on sample concentration, which both yield accurate (2 to 0.02%) molecular weight information on intact proteins without any additional sample preparation, electroblotting, destaining, etc. The remaining 97% of a fraction can then be used for other analyses, such as Edman sequencing, amino acid analysis, or proteolytic digestion and sequencing by tandem mass spectrome-

try. This 2D HPLC protein purification and identification system was used to isolate the src homology (SH2) domain of the nonreceptor tyrosine kinase pp60src and β-lactamase, both inserted into E. coli, as well as a number of native proteins comprising a small portion of the E. coli proteome. © 1998 Academic Press

A two-dimensional (2D)2 separation using isoelectric focusing (IEF) followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) is a powerful method for separating the thousands of proteins found in a total cell lysate. These 2D gels have also become highly sensitive and routine in the 22 years since first reported by O’Farrell (1). IEF–SDS–PAGE is not without its limitations, and it was to circumvent these limitations that work was begun on an alternative approach to analyzing crude cell lysates. Casting, loading, and running a gel is a labor-intensive process where slight differences in migration are usually overcome by casting large lots of gels simultaneously and running them immediately in parallel (2). There is also a concern that free, unpolymerized acrylamide can bind to the amino-terminal of a protein, blocking Edman degradation chemistry (3). Detection of proteins in gels, using methods such as silver staining (4–6) and autoradiography (1), is slow compared to UV detection and can require growing the cells in an isotopically enriched broth. Once a 2D gel is run, a pro-

\[ \text{Equation} \]

1 To whom correspondence and reprint requests should be addressed. Fax: 919-962-1381.

2 Abbreviations used: 2D, two-dimensional; IEF, isoelectric focusing; MALDI-TOF/MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; ESI/MS, electrospray mass spectrometry; RPLC, reversed-phase liquid chromatography; SEC, size-exclusion liquid chromatography; BSA, bovine serum albumin; TFA, trifluoroacetic acid.
cess which can take days (7), additional information is always required to determine the identity or modifications of the underlying proteins (8), since the isoelectric point and molecular weight information produced by SDS–PAGE are accurate to only about ±10% (3). Gel spots can be analyzed for molecular weight by electrophoresis followed by mass spectrometric analysis (8, 9) or blotting to a membrane (10–12) for MALDI-TOF/MS of the intact protein. Gel spots can also be electroeluted into free solution for digestion and the fragments (13, 14) chemically sequenced. Spots have also been electroblotted onto a membrane for sequencing (15–18). Still other routes for identification of gel isolated proteins have used either RPLC–ESI/MS for peptide mapping (19) or spots cut from the gel which can be digested in situ with tandem MS performed on the resultant peptides (4, 20–23).

In other words, while gels provide extremely high resolution separations of proteins, there are sufficient drawbacks to them which makes it important to explore alternative separation schemes which do not modify the proteins and promise to be faster and more reproducible. One alternative to separating the mixtures of proteins in total cell lysates is to use high-performance liquid chromatography (HPLC). Since no acrylamide is present in HPLC, the possibility of blocking the amino-terminals of the proteins is reduced. HPLC also provides a high level of reproducibility because of the easy availability of high-quality commercial columns which can be purchased, stored indefinitely, and used repeatedly. Most HPLC systems are easily automated and by utilizing fraction collectors and autosamplers, these systems can operate unattended. The preferred method of detection in HPLC is by UV absorbance of the peptide bond because of its speed, sensitivity, and linearity over three orders of magnitude (24). Following HPLC separation, the fractions are left in free solution and are immediately available for further analysis. Unfortunately, while overcoming many of the disadvantages of 2D gels, HPLC alone simply does not yet provide enough chromatographic resolution (25–29) to adequately separate an entire cell lysate.

However, a comprehensive 2D HPLC system described in this report shows dramatically increased resolution over standard “one-dimensional” HPLC in the way a 2D gel is a dramatic improvement over performing a one-dimensional gel experiment. For two components to remain overlapped at the end of a IEF–SDS–PAGE experiment; they must both have almost the same isoelectric points and molecular weights. For two components to coelute from this 2D HPLC system, they must both have almost the same molecular weights as well as hydrophobicities; therefore the success of separating the hundreds or thousands of proteins expressed in various cell lysates increases dramatically (30, 31). Furthermore, the 2D HPLC system in this report distinguishes itself from previous 2D HPLC systems (32) for a number of reasons. It subjects all of the soluble proteins in the lysate to both dimensions of separation, thus it is “comprehensive,” so the two-dimensional chromatogram contains data on an organism’s entire soluble proteome, rather than just on a few proteins as happens with a heart-cutting system (33–43). The volume of the fraction transferred to the second dimension is usually half the volume of the sharpest peak. Proteins which begin to separate by size, but are not baseline resolved when they elute from the first dimension, are thus transferred as separate peaks using this system rather than being recombined during the transfer process; even if two proteins have exactly the same hydrophobicities, the partial separation performed by the first dimension is not lost (30). Thus, maintaining almost all of the separation obtained by the first dimension and not using stopped flow techniques (44–49) maximizes this 2D HPLC system’s chromatographic resolution. Still other comprehensive 2D systems for the analysis of proteins (50–52) have lacked the resolution, ruggedness, and preparative capacity of this system.

Once proteins from cell lysates are chromatographically separated with this 2D HPLC system, they can be detected at 215 nm regardless of the UV absorbance of the buffers used in the size-exclusion mobile phase because of the desalting effect of the RPLC in the second dimension. Following UV detection, a fraction collector using six deep-well polypropylene microtiter plates can provide room for 576 individual fractions, which are immediately available for matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF/MS), electrospray ionization mass spectrometry (ESI/MS), Edman sequencing, or other analytical techniques. Chromatographic separation and UV detection, from injection to collection of the final fraction requires 6.5 h. Adding 30 min for sample preparation prior to this, 30 min for analysis by MALDI-TOF/MS and ESI/MS following the separation, and 30 min to load a sequencer cartridge means that a protein of interest can progress from the stage of being present in Escherichia coli cells suspended in their fermentation broth to being loaded on the Edman sequencer ready for overnight analysis, in a single 8-h day.

This is a report on the design and implementation of a 2D HPLC system which uses size-exclusion chromatography followed by reversed-phase liquid chromatography for the high-resolution chromatographic separation of proteins. These separations are followed by both MALDI-TOF/MS and ESI/MS for higher resolution molecular weight analyses, and amino-terminus Edman sequencing. This system was used to isolate and identify two nonnative proteins expressed in E. coli: the src
homology (SH2) domain of pp60c-src and β-lactamase, as well as a number of native E. coli proteins.

**EXPERIMENTAL**

Size-exclusion chromatography. The first dimension of SEC consisted of eight 7.8-mm i.d. × 300-mm columns packed with 5 μm diameter material having a pore size of 250 Å (G3000SWXL, Toso-Haas, Montgomeryville, PA) placed in series and connected by 10-cm lengths of 126-μm i.d. × 1.58-mm o.d. PEEK tubing (Upchurch Scientific, Oak Harbor, WA). This was changed for the final proteome mapping experiment, in order to maximize chromatographic efficiency but while remaining within the operating pressure of the HPLC pump, by reducing the number of G3000SWXL columns from eight to six, and adding six G4000SWXL (Toso-Haas) columns which used 8-μm-diameter material having a pore size of 450 Å but otherwise identical dimensions. This change increased the overall effective column length from 2.4 to 3.6 m.

Two different mobile phases were used, either the nondenaturing mobile phase which consisted of 100 mM sodium chloride (Sigma, St. Louis, MO), 10 mM Tris–HCl (International Biotechnologies, New Haven, CT), 1 mM ethylenediaminetetraacetic acid (EM Science, Gibbstown, NJ), pH 7.5, adjusted with potassium hydroxide (J. T. Baker, Phillipsburg, NJ); or the "denaturing" mobile phase which consisted of 50 mM potassium phosphate (Sigma), 3 M urea (Sigma), pH 7.1, adjusted with phosphoric acid (EM Science). The mobile phase was filtered through a 0.2-μm nylon membrane (Gelman Sciences, Ann Arbor MI). Flow rates were 250 μL/min for the 8-column system and 150 μL/min for the 12-column system. The exclusion volume of each column was 5.6 mL, and the inclusion volume 11.2 mL, resulting in system exclusion and inclusion volumes of 44.8 mL (8 × 5.6 mL) and 89.6 mL (8 × 11.2 mL), respectively. The mobile phase was supplied by a Hewlett-Packard (Model 1050, Palo Alto, CA) pump and controller, labeled “Pump 1” in Fig. 1.

Reversed-phase chromatography. The second dimension consisted of two polystyrene divinylbenzene 2.1-mm i.d. columns (R2/H, PerSeptive Biosystems, Framingham, MA) placed in parallel. All experiments used 33-mm-long columns except the final proteome mapping experiment, which used 10-cm-long columns. In all experiments the same steep gradient at 1.5 mL/min. was pumped by a HP 1090, labeled “Pump 2” in Fig. 1. Mobile phase A was water with 0.1% trifluoroacetic acid (Pierce, Rockford, IL) and phase B was acetonitrile (Burdick and Jackson, Muskegon, MI) with 0.1% TFA, both passed through 0.2-μm nylon membrane filters (Gelman Sciences). Gradients started at 15% B, then at 0.1 min to 27% B, at 3.3 min 55% B, and then back to 15% B for a 0.7-min reequilibration (total time between injections, 4 min). The SEC column outlet was connected to a 4-port valve (Valco Instruments, Houston, TX) which alternated the connection between the two RPLC columns, thus serving as an injector.

UV detection and fraction collection. The RPLC column outlets were connected to a second 4-port valve. This 4-port valve switched in tandem with the first valve. The SEC mobile phase flowed to waste, after the proteins that it initially carried were trapped on the head of the RPLC column. The UV detector (Series 1050, Hewlett-Packard) was equipped with a standard analytical flow cell and monitored absorbance of the peptide bond at 215 nm at a data collection rate of 2 Hz. The outlet of the UV detector was connected to a fraction collector (222 XL, Gilson, Middleton, WI) via 1 m of

**FIG. 1.** Schematic of the 2D LC/LCMS system showing the 8 SEC columns arranged in series for an effective column length of 2.4 m, the parallel 33-mm-long RPLC columns in the LC/LC interface, and the location of the UV detector prior to fraction collection.
126-μm i.d. × 1.58-mm o.d. PEEK tubing (Upchurch Scientific). The fraction collector used a custom collection routine supplied by the manufacturer. The routine allows the user to set a delay time before beginning collection, which was 45 s, and a collection time per vial, which was 10 s, for a volume of 250 μL. The fractions were collected in six deep-well polypropylene microtiter plates (Beckman, Fullerton, CA) which correspond in space to the 2D chromatograms generated by the UV detector.

Mass spectrometry and Edman sequencing. MALDI-TOF/MS analysis of interesting fractions was conducted using a bench-top instrument (LaserMat, FinniganMat, San Jose, CA) with sinapinic acid (Sigma) as the matrix. Higher resolution molecular weights were obtained using a ESI/MS on a single quadrupole mass spectrometer (API 100, PEScience, Norwalk, CT). Fractions were infused from a glass syringe at 1 μL/min (10 μL, Hamilton, Reno, NV) using a syringe pump (11, Harvard Apparatus, South Natick, MA) without any sample pre-treatment. The amino terminus sequence of each protein was determined using an automated Edman degradation protein sequencing system (G1005A, Hewlett-Packard).

Timing and data handling. The chromatographic system was controlled by a Macintosh computer (IIfx, Apple Computer, Cupertino, CA) running a custom program written in-house with LabView (3.0, National Instruments, Austin, TX) software and using a data acquisition board (NB-MIO-16XL, National Instruments) for switching the valves, communicating with Pump 2, and collecting the UV data. Data from the UV detector were background subtracted to remove refractive index changes detected during the mobile-phase gradient and displayed using Spyglass Transform 3.3 (Fortner Research, Sterling, VA). MALDI-TOF/MS and ESI/MS data were displayed using the control software supplied with the instruments.

Samples. E. coli strain DH5α was transformed with a plasmid expressing SH2 and β-lactamase. A single colony from an LB-agar plate (Difco Laborato-
ries, Fisher Scientific, Pittsburgh, PA) was used to inoculate 5 mL of cells which were allowed to grow overnight to saturation (≈10⁸ cells/mL). The cells were centrifuged at 7000 g for 10 min at 4°C (RC-50, Sorvall Instruments, Newtown, CT). The medium was decanted and the cells were resuspended in 1 mL of the SEC mobile phase, either denaturing or nondenaturing, depending on that particular analysis, and sonicated for 3 min (450, Branson Ultrasonics, Danbury, CT) to rupture the cells. The sample was centrifuged to pellet the particulate again at 7000g for 10 min. The supernatant was then withdrawn and transferred to another tube for storage. This supernatant could then either be frozen at −80°C or diluted with additional SEC mobile phase or immediately injected, depending upon application. A typical injection consisted of 100 μL of the supernatant being loaded onto the first dimension. Bovine serum albumin for the recovery experiment (Sigma) was used as received and dissolved in denaturing SEC mobile phase.

RESULTS AND DISCUSSION

Each comprehensive two-dimensional analysis using this system, a schematic of which is shown in Fig. 1, was performed as follows. A single 100-μL injection was made onto the first of eight SEC columns connected in series. As the analytes would elute from these columns, they were alternately loaded in 4-min wide plugs onto either of a pair of RPLC columns. If an SEC peak contained multiple components of similar size, but different hydrophobicities, then the components would be separated by the reversed-phase column. Otherwise, components of similar hydrophobicity eluted from the SEC columns, and one of the RPLC columns would simply sharpen the peaks prior to detection. After eluting from the reversed-phase column, the analytes passed through a UV detector set to 215 nm and then to the fraction collector.

The LC/LC interface, as shown in Fig. 2, served as an “injector” for the second dimension. Since no proteins could elute from the SEC columns until one exclusion volume passed, the effluent was diverted to waste for the first 150 min following an injection. Then the SEC effluent was manually connected to the first 4-port valve so effluent would flow through the upper RPLC column after it had been equilibrated with 15% B from the gradient pump. Since the SEC mobile phase lacked an organic component such as acetonitrile or methanol,
proteins eluting from the SEC were trapped on the upper RPLC column while the SEC mobile phase containing salts, buffers, etc. flowed to a waste container. This loading continued for 4 min at which point the 4-port valve switched, connecting the sample laden upper RPLC column to the gradient pump and allowing any proteins then eluting from the SEC columns to be trapped on the lower RPLC column. Simultaneously, the upper RPLC column was run through a gradient of increasing organic concentration while the lower RPLC column was being loaded for 4 min. The valves switched every 4 min for the next 240 min of the SEC columns' elution window.

As was expected with an HPLC-based system, reproducibility was excellent. Whereas one-dimensional HPLC systems describe peaks with one data point, their retention time, this system uses a two data points, SEC and RPLC retention times, in order to describe the location of a single peak. Five peaks resulting from naturally occurring proteins in E. coli were monitored during four experiments over a period of 5 days. All HPLC conditions were held constant, and the cell lysate was stored at -20°C between experiments. The average standard deviation in SEC retention time was 1.06 min (0.34%) and the average standard deviation in RPLC retention time was 1.03 s (0.43%). This level of reproducibility made the 2D chromatograms indistinguishable to the naked eye and suggested that this system could be used to monitor the regulation of native protein expression.

When designing this system, it was very important to choose a chromatographic mode in the second dimension amenable to mass spectrometry. It was also necessary that the second dimension of chromatography could perform a separation in a time less than half that of the width of a peak eluting from the first dimension, in order to maintain the separation affected by the first dimension (30, 53, 54). Reversed-phase liquid chromatography was chosen because of the low pH and complete volatility of the mobile phase, which simplified the transfer of fractions to MALDI-TOF/MS and ESI/MS, techniques which would have been detrimentally affected by the presence of nonvolatile buffers. Furthermore, there are many rugged, reliable, and efficient columns commercially available for RPLC, many of which are geared toward providing fast separations. Finding an orthogonal first dimension, one which would separate the proteins on the basis of some property other than hydrophobicity, became necessary; otherwise, no advantage would result from performing

**FIG. 5.** (A) Electrospray mass spectrum of fraction indicated in Fig. 3 and analyzed in Fig. 4. The various charge states of SH2 present are labeled. (B) Plot of the reconstructed molecular weight of SH2 which would generate the ions displayed in A. This reconstructed molecular weight of 12,290 Da is within 0.03% of the expected molecular weight of SH2 (12,287 Da).
a prior dimension of separation. Size exclusion chromatography was chosen so an approximate estimate of a protein’s molecular weight could be determined prior to any off-line analysis. SEC was also compatible with the high salt and buffer concentrations generated during the fermentation process which would normally cause a problem for a technique such as gradient ion exchange chromatography. Unfortunately, SEC is generally thought of as a low-resolution chromatographic technique, although this is not necessarily true. While it does not allow the chemist to utilize gradient elution to improve the separation, multiple columns can be linked together in series which increases the effective column length thus improving chromatographic efficiency. In this case, eight SEC columns were placed in series in order to generate ~90,000 theoretical plates. While the same 2.4 m of effective total column length could have been obtained at reduced cost by making use of packed capillary chromatography, it was determined that submillimeter bore columns lacked the capacity necessary to load on sufficient sample if a usable amount (10–20 pmol/component) of proteins was to be returned (50–52).

The entire 2D HPLC system was tested for recovery using a protein standard. An injection containing 17 pmol of bovine serum albumin (BSA) was made onto the system, as determined by amino acid analysis of the solution from which the injection was made. The denaturing mobile phase was used and all experimental conditions were identical to those of standard 2D experiments separating cell lysates. The BSA eluted in two RPLC analyses, as would be expected in a comprehensive 2D HPLC system, and the corresponding fractions were submitted for amino acid analysis. Results in triplicate showed the presence of 7 pmol of bovine serum albumin in each well, for a recovery of 80%. The sample loss could have resulted from protein adsorbing to the walls of the fraction collector, sample handing losses when extracting the fractions from the wells, or binding to the columns. No BSA was detected by UV or MS analyses during subsequent blank injections. This level of recovery, though, could not be expected for every protein found in a cell lysate, but is

FIG. 6. A 2D UV chromatogram of an E. coli lysate grown to overexpress β-lactamase. The fraction indicated at 302 min of the SEC and 130 s of the RPLC was collected and subjected to further analysis.
instead supplied to give an estimate for a large protein analyzed under normal operating parameters. Unfortunately, as in all HPLC systems utilizing reversed-phase chromatography (55), very hydrophobic membrane proteins, heavily glycosylated proteins, or very large (>100 kDa) proteins would be expected to show a low level of recovery and poorer chromatographic separation. If the analysis of these types of proteins were to be routinely performed on such a 2D HPLC system, the second dimension of RPLC could be replaced with gradient ion exchange chromatography or hydrophobic interaction chromatography. These techniques, though, would require an additional sample cleanup step prior to mass spectrometry.

A 2D HPLC chromatogram of the separation of an E. coli lysate inoculated with a plasmid containing the gene for the SH2 domain of pp60c-src as detected at 215 nm is shown in Fig. 3. The 2D chromatogram is the result if each of the RPLC chromatograms is stacked side by side and looked upon from a top-down perspective. A peak’s SEC retention time can be obtained from the Y axis, while its RPLC retention time is depicted on the X axis. The peaks were evenly spread out, using most of the available peak area, though a trend can be observed in the data of larger proteins showing greater retention time in the RPLC dimension than do smaller proteins. Another aliquot of the same bacterial cell sample was analyzed by packed capillary LC/MS and the data reported elsewhere (56); those data showed that a one-dimensional HPLC system was capable of chromatographically resolving only ≈50 peaks, whereas the 2D HPLC system reported here resolved at least ≈250 peaks. The overexpressed protein of interest in this sample SH2 has a molecular weight of 12287 Da, so it was expected to elute fairly late in the SEC dimension. In fact, the spot at 325 min on the Y axis (SEC) and 50 s on the X axis (RPLC) was immediately assumed to be that of SH2 because of its increased intensity relative to that of native proteins and its SEC retention time. Two microliters of the corresponding fraction, which had a total volume of 250 μL, was removed and applied to a target for MALDI-TOF/MS analysis, the spectrum from which is shown in Fig. 4. The apparent molecular weight of 12,255 Da was within 0.3% of the expected molecular weight, but in order to ascertain the spot’s identity another 5 μL was removed from the fraction for an electrospray infusion experiment, the spectrum of which is shown in Fig. 5A along with the reconstructed molecular weight in Fig. 5B. In the ESI/MS experiment, the observed molecular weight of 12,290 Da fell within 0.03% of the expected mass so the remaining ≈240 μL of the fraction was submitted for Edman sequencing. The sequencer detected 300 pmol in the first cycle, and five more cycles were carried out to verify that the amino-terminus sequence correctly corresponded to the expected sequence of MD-SIQP. This 2D HPLC analysis of SH2 was of particular

![Fig. 7. Corresponding off-line MALDI-TOF/MS mass spectrum resulting from the same peak indicated in Fig. 6, showing a base peak of 29,015 m/z along with the doubly protonated species at 14,483 m/z. The base peak is within 0.4% of the expected molecular weight of β-lactamase (28,907 Da).](image)
interest because SH2 is a basic protein, having an isoelectric point of 9.1. Such basic proteins are notorious for poor performance (57) in the IEF dimension of 2D gels; however, SH2 presented no particular problem for this HPLC system.

In order to further evaluate the system with a larger overexpressed protein of interest, E. coli was grown which had been inoculated with a plasmid containing the gene for \(\beta\)-lactamase in order to confer ampicillin resistance. This protein has a \(pI\) of 6.4 and a molecular weight of 28,907 Da. A 2D HPLC chromatogram from an injection of 100 \(\mu\)L of cell lysate supernatant (\(\sim 1\) mg/mL total protein concentration) run under nondenaturing conditions in the first dimension and detected at 215 nm is shown in Fig. 6. Using the nondenaturing mobile phase, the spots are shifted to appreciably longer SEC retention times, which results from the smaller hydrodynamic diameters of the folded proteins, and greater inclusion into the pores of the SEC stationary phase. Upon first observing this 2D chromatogram, all four of the most intense spots were close enough in SEC retention time to have been \(\beta\)-lactamase. All four were analyzed by MALDI-TOF/MS and ESI/MS, but only the spot at 302 min of the SEC and 130 s of the RPLC was found to show spectra within 1% of 28.9 kDa. As seen in Fig. 7, the MALDI-TOF/MS spectrum shows the intact protein at \(m/z\) 29,015, which is within 0.4% of the expected value. The peak at \(m/z\) 14,483 results from the doubly protonated species. The ESI/MS spectrum is seen in Fig. 8A. Reconstructing the intact mass of the protein yields a 28,917 Da as seen in Fig. 8B. This is within 0.03% of the expected mass. The corresponding fraction was submitted for Edman sequencing, where only a single species was detected, and its amino terminus correctly matched that expected of \(\beta\)-lactamase: HPETLV. During this Edman degradation experiment, only 20 pmol was present in the first cycle, which was a much lower level than observed during the SH2 analysis. This was not an unexpected result owing to the lower level of expression in E. coli and the factor of 5 reduction in sample load; these data were also supported by a proportional loss in the signal to noise ratio of the ESI/MS experiment.

In each of the 2D LC/LC experiments shown thus far, there were \(\sim 250\) spots visible with UV detector.

**FIG. 8.** (A) Electrospray mass spectrum of fraction indicated in Fig. 6 and analyzed in Fig. 7. The various charge states of the \(\beta\)-lactamase present are labeled. (B) Plot of the reconstructed molecular weight of \(\beta\)-lactamase which would generate the ions displayed in A. The reconstructed molecular weight of 28,917 Da is within 0.03% of the expected molecular weight of \(\beta\)-lactamase (28,907 Da).
The maximum number of spots that this system could have theoretically resolved under optimal conditions was approximately 800. This value was arrived at by multiplying the maximum number of peaks capable of being resolved in a single RPLC separation by the number of RPLC separations which analyzed the SEC elution window. An average RPLC peak was 10 s wide, so 20 peaks could fit into the ≈200 s of usable RPLC elution time. The ≈160-min-wide SEC elution window was broken up into 40 RPLC analyses. Therefore, multiplying 20 (the maximum number of RPLC peaks per analysis) by 40 (the maximum number of RPLC analyses) yields a theoretical limit of 800 spots. The two major reasons why only ≈30% of this was observed are as follows. First, the theoretical calculation assumes that all the components will be randomly distributed, but this was not the case primarily because the larger proteins to exhibited greater retention on the RPLC columns than their smaller counterparts. This was probably a sieving effect of the “through-pores” in the RPLC stationary phase (58) or a tendency for larger proteins in E. coli to be more hydrophobic on average. Second, the calculation assumes that all spots will fall above the limit of detection; clearly this was not the case, though it could be improved by solubilizing a greater fraction of the proteome in both dimensions and using a more sensitive detector. The maximum protein load which this 2D HPLC system could handle before a perceived qualitative loss in chromatographic resolution was 2 mg, although most 2D experiments were typically conducted with 0.2 mg total weight of protein.

For some rudimentary proteome mapping experiments (59–64), this comprehensive LC/LC system was used in a slightly elongated configuration to increase its resolving power. The separation shown in Fig. 9 was obtained by increasing the effective column length of the first dimension by 50%, reducing the SEC flow rate by 40%, and tripling the length of the second-dimension columns. Not only does this increase the number of theoretical plates in the first

**FIG. 9.** A 2D UV chromatogram of an E. coli lysate showing a selection of native proteins isolated using an elongated 2D HPLC system, where the first dimension has been lengthened from 2.4 m of effective total column length to 3.6 m, the flow rate slowed from 250 to 150 μL/min, and the length of the RPLC columns tripled to 10 cm.
The data presented here showed that this 2D HPLC system was useful for the quick isolation of overexpressed proteins of interest from bacterial cell lysates. The system could sit for long periods of time and then be quickly activated when needed. It was completely automated, requiring little skill or experience to operate. It showed good protein recovery, adequate resolution, and sensitivity equivalent to the Edman sequencing technique which followed it. Unfortunately for proteome mapping efforts, this 2D HPLC system in its present form lacked the resolution of 2D gels. However, it could probably provide adequate resolution for the mapping of small proteomes (23, 64) or mapping portions of larger proteomes. The 2D HPLC system was faster than gels (cells to isolated protein in 1 day) and of course did not block the amino termini of the proteins with free acrylamide. A number of methods were suggested to improve this system should it be used strictly for proteome mapping, such as further lengthening the first dimension, changing from SEC to gradient ion exchange HPLC, and replacing the RPLC in the second dimension with hydrophobic interaction chromatography for improved performance of hydrophobic, large (100 kDa), or glycosylated proteins. This 2D HPLC system returned rather large quantities (~1–10 pmol) of protein in a convenient solution format of water/acetonitrile at low pH; and though this would not be conducive to activity assays, it made the many follow-up identification techniques, e.g., MALDI-TOF/MS, ESI/MS, Edman sequencing, Nanospray/MS, very straightforward. The data in Fig. 9/Table 1 also suggested that for an organism such as E. coli, an experiment could be performed where a total E. coli lysate could be separated into ~400 fractions of ~10 proteins each in a short amount of time. From there an autosampler could easily withdraw the fractions and pass them through an immobilized enzyme column (65) for proteolysis. The resultant peptides could then be sequenced by tandem mass spectrometry techniques, and the precursor proteins could be identified using a database searching program (66).

ACKNOWLEDGMENTS

The authors appreciate the help of Will Burkhart and Mary Moyer for Edman sequencing; Kevin Blackburn for MS/MS sequencing; Doug Sheeley for consultation; and Ramani Narayanasami, Bruce Wisely, and Martin Rink for the E. coli samples. All are employees

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Protein</th>
<th>Mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02349</td>
<td>30S ribosomal protein S1</td>
<td>61,159</td>
</tr>
<tr>
<td>P22257</td>
<td>Trigger factor</td>
<td>48,023</td>
</tr>
<tr>
<td>P06994</td>
<td>Malate dehydrogenase</td>
<td>36,511</td>
</tr>
<tr>
<td>P26427</td>
<td>Alkyl hydroperoxide reductase C22</td>
<td>20,630</td>
</tr>
<tr>
<td>P02408</td>
<td>50S ribosomal protein L10</td>
<td>17,580</td>
</tr>
<tr>
<td>P02356*</td>
<td>30S ribosomal protein S5*</td>
<td>17,472</td>
</tr>
<tr>
<td>P02358</td>
<td>30S ribosomal protein S6</td>
<td>15,703</td>
</tr>
<tr>
<td>P02392</td>
<td>50S ribosomal protein L7/L12</td>
<td>12,164</td>
</tr>
<tr>
<td>P26604</td>
<td>HDEA precursor</td>
<td>11,858</td>
</tr>
<tr>
<td>P02901*</td>
<td>Acyl carrier protein*</td>
<td>8,508</td>
</tr>
<tr>
<td>P36996</td>
<td>Cold shock-like protein cspc</td>
<td>7,271</td>
</tr>
</tbody>
</table>
REFERENCES

TWO-DIMENSIONAL CHROMATOGRAPHIC SEPARATION OF PROTEINS