ELSEVIER

Journal of Chromatography B, 686 (1996) 165–175

JOURNAL OF CHROMATOGRAPHY B: BIOMEDICAL APPLICATIONS

Internal sequence analysis of proteins eluted from polyacrylamide gels

Byung Joon Hwang[a,b], Alan J. Smith[c], Gilbert Chu[a,b,*]

[a]Department of Medicine, Stanford University Medical Center, Stanford, CA 94305, USA
[b]Department of Biochemistry, Stanford University Medical Center, Stanford, CA 94305, USA
[c]Protein and Nucleic Acid Facility, Stanford University Medical Center, Stanford, CA 94305, USA

Received 21 February 1996; revised 20 May 1996; accepted 20 May 1996

Abstract

We have developed an elution–digestion–sequencing (EDS) method, which yields the internal amino acid sequence of partially purified proteins. The overall yield for the method was greater than 60%. The method yielded peptide peaks that could be sequenced on HPLC for all tested proteins with masses from 45 to $200 \cdot 10^3$ and yielded internal amino acid sequence information when as little as 10 pmol of partially purified protein was used as the starting material. The EDS method was extremely reliable and gave sequence information for each of 25 proteins tested, including high-molecular-mass proteins ($M_r > 100 \cdot 10^3$) that were difficult to sequence by other methods.

Keywords: Elution–digestion–sequencing method; Protein sequencing; Amino acid sequencing; Internal amino acid sequencing; Proteins

1. Introduction

Protein sequencing is of critical importance in biological research. A partial amino acid sequence may be used either to identify a protein with a previously described protein or to design oligonucleotide probes for cloning the corresponding gene. Sequencing methods for internal amino acid are particularly valuable because many proteins are modified at the N-terminus either in vivo or during purification, so that they are not susceptible to Edman degradation [1]. Furthermore, more sequence information can be obtained from the same quantity of protein, facilitating more precise identification of the protein and the design of more effective oligonucleotide probes for gene cloning.

Because proteins of biological interest may often be of low abundance, highly efficient methods are needed for sequencing proteins in small amounts at low concentrations. Such methods should require only partial purification of the target protein, since most of the protein may be lost during the final purification steps. Currently, the most powerful methods combine partial chromatographic purification with the high resolving power of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). The electrophoretically resolved proteins are bound either to nitrocellulose as described by Aebersold et al. [2] or to polyvinylidene difluoride (PVDF) as described by Bauw et al. [3] and then enzymatically digested. The methods are ca-

*Corresponding author.

0378-4347/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
PII S0378-4347(96)00230-7

pable of analyzing small amounts of protein [4] and consist of: resolution of the partially purified protein fraction by SDS–PAGE; electrophoretic transfer from the gel matrix to a nitrocellulose- or PVDF membrane; detection of protein bands by staining with Ponceau S or amido black; enzymatic cleavage of individual proteins in situ on the membrane; elution of the peptide fragments from the membrane; separation of the peptides by reversed-phase high-performance liquid chromatography (HPLC) and microsequencing of the peptides.

Although this method has been widely used, there are a number of difficulties that may interfere with its successful application for certain proteins. For example, high-molecular-mass proteins may be transferred inefficiently to the membrane. Conversely, low-molecular-mass proteins may be washed off the membrane prematurely during subsequent steps in the procedure. The proteins are detected on the membrane by staining with Ponceau S or amido black, which are significantly less sensitive than Coomassie blue [5]. In situ cleavage of the protein may be inefficient because of adsorption of the protease to the membrane, even when the membrane is pretreated with polyvinyl-pyrrolidone 40. HPLC peaks for peptide fragments may be decreased or lost because of poor elution from the membrane, poor enzymatic cleavage, or inaccessibility of some cleavage sites when the protein is immobilized on the membrane.

We developed a new method for obtaining protein sequence information after encountering difficulties in sequencing the protein XPE-BF (xeroderma pigmentosum group E binding factor) [6]. XPE-BF binds to damaged DNA and is deficient in a subset of patients from complementation group E of xeroderma pigmentosuma. It presented two common problems for protein microsequencing, i.e., low abundance and high molecular mass ($125\cdot10^3$). Furthermore, purification to homogeneity resulted in a large loss of yield during the final chromatographic step.

The new procedure for obtaining sequence information is suitable for partially purified proteins of both high- and low molecular masses. The elution–digestion–sequencing (EDS) method consists of: elution from the gel facilitated by SDS and digestion with Lys-C protease; heptane–isoamyl alcohol extraction to remove the SDS; re-digestion of the protein in a solution with trypsin or Lys-C protease; and sequencing of the peptides.

The method has several advantages. It avoids transfer of proteins to a solid support. Elution from the gel is robust enough to accommodate both high- and low-molecular-mass proteins. Staining with Coomassie blue is more sensitive than with Ponceau S or amido black. The final proteolysis is carried out in solution and is therefore more efficient than digestion on a solid support.

2. Experimental

2.1. SDS–PAGE

Proteins were resolved by electrophoresis in a stacking gel of 3% polyacrylamide (acrylamide–bisacrylamide at a ratio of 29:1) with 50 V and then in a resolving gel of 7.5% or 12.5% polyacrylamide with 120 V for 60 min at room temperature in either a 0.75 or 1.5 mm thick gel using a mini-gel apparatus (Hoefer, San Francisco, CA, USA) [7]. Gels containing proteins for digestion were stained in 0.1% Coomassie blue, 50% methanol and 5% acetic acid for 5–10 min, then exhaustively destained in 50% methanol and 10% acetic acid. The test proteins, myosin, β-galactosidase, phosphorylase *b*, bovine serum albumin (BSA), ovalbumin and RNAse A, were obtained commercially (Bio-Rad, Richmond, CA, USA). XPE-BF was purified as described previously [6]. The protein concentrations were determined by amino acid analysis [8].

2.2. Protein elution from SDS–PAGE gel

The protein bands were cut out of the gel and washed in water to remove residual acetic acid and gel buffer components that might interfere with subsequent proteolysis. The gel slice was then diced into small chips, which were dehydrated in a Speed-Vac. The protein was eluted from the gel chips by incubation in buffer A containing SDS and Lys-C protease (1 mg/ml, stock solution, sequencing grade, Boehringer Mannheim, Indianapolis, IN, USA). (See Table 1 for details).

Table 1
Protocol for the EDS method

Elution of protein from SDS–PAGE gel
1. Resolve proteins by SDS–PAGE and stain the gel with Coomassie blue.
2. Excise the desired protein band with a razor blade, wash the gel slice four times in 15 ml of water for 15 min at room temperature, dice into small chips (1 mm square), and completely dry the gel chips using a SpeedVac for at least 1 h.
3. Digest the protein in situ by barely covering the chips with buffer A (20 m*M* Tris–HCl, pH 8.0) containing both 0.04% SDS and Lys-C protease at a concentration of 1% (w/v). As the gel chips swell, add buffer A containing 0.04% SDS to keep them covered. Spot a small aliquot onto pH paper to verify that the sample has been buffered to pH 8.0. Elute the protein fragments by gently rocking the sample overnight at 37°C. Save the eluate.
4. Elute residual protein fragments twice more with water for 3 h at 37°C and once with 100 μl of acetonitrile for 1 h at 37°C with gentle rocking. Save the eluates.
5. Combine the eluates from steps 3 and 4, and dry in a SpeedVac.

Removal of SDS by heptane–isoamyl alcohol extraction

6. Resuspend the dried sample in 100 μl of water and acidify by adding 5 μl of TFA.
7. Extract the acidified sample with 100 μl of heptane–isoamyl alcohol (4:1, v/v) by vigorous vortex-mixing for 5 min. Centrifuge the mixture for 1 min at room temperature and discard the upper organic phase. (The interphase contains Coomassie blue and the lower phase contains the peptides). Re-extract the interphase and the lower phase with heptane–isoamyl alcohol.
8. Dry the remaining interphase and lower phase using a SpeedVac (1–2 h).
9. Resuspend the pellet in 100 μl of acetonitrile and dry using a SpeedVac.
10. Resuspend the pellet in 100 μl of water and dry using a SpeedVac.

Digestion of protein in solution and sequencing of peptides

11. Resuspend the sample and digest to completion with either Lys-C or trypsin.
12. Separate the peptides by HPLC and sequence by Edman degradation.

2.3. SDS removal by heptane–isoamyl alcohol extraction

SDS remaining in the eluate was removed by extraction with heptane–isoamyl alcohol following the method of Bosserhoff et al. [9]. The interphase and lower phase were dried in a SpeedVac, and the sample was resuspended in acetonitrile and dried, then resuspended in water and dried again. (See Table 1 for details).

2.4. SDS removal by AG11A8 ion-exchange chromatography

The column was prepared by packing a 200-μl Pipetman tip with 100 μl of AG11A8 resin (Bio-Rad, Richmond, CA, USA) and washing with 0.1% trifluoroacetic acid (TFA) as described by Kapp and Vinogradov [10]. The column was loaded with protein and washed with 0.1% TFA. The flow-through and wash fractions were pooled and dried in a SpeedVac. Recovery was quantitated by comparing the recovered sample to serial dilutions of the starting material after resolution by SDS–PAGE and staining the gel. This method of removing SDS was not part of the final procedure (see Section 3).

2.5. Trypsin digestion

Trypsin digestion was performed as described by Stone et al. [11]. The dried sample was resuspended in 50 μl of 8 *M* urea, 50 m*M* Tris–HCl, pH 8.0. The protein was then reduced by addition of 0.5 μl of 1 *M* DTT followed by incubation for 15 min at 50°C. The denatured, reduced sample was cooled to room temperature and alkylated by adding 1 μl of 500 m*M* iodoacetamide and incubating for 15 min at room temperature in the dark. The volume was then increased to 200 μl with the addition of water to dilute the urea and salt. Trypsin (sequencing grade, Boehringer Mannheim) was added to 10% (w/w) and the mixture was incubated for 10–18 h at 37°C with gentle rocking.

2.6. HPLC separation of tryptic peptides

To remove residual polyacrylamide fragments prior to HPLC, a 200-μl Pipetman tip was packed

with 20 μl of siliconized glass wool and washed with 300 μl acetonitrile. The sample containing tryptic peptides was loaded and the flow-through was collected. Residual peptides on the column were eluted with 100 μl of acetonitrile. The pooled volume was evaporated down to 25 μl, then diluted up to 100 μl with 0.1% TFA prior to HPLC analysis. Peptides were separated on a reversed-phase column (220×2.1 mm I.D., Aquapore 300, Applied Biosystems, Foster City, CA, USA) using a solvent delivery system (Model 140B, Applied Biosystems) and a programmable UV–Vis detector (Model 785A, Applied Biosystems). Peptides were injected in a volume of 100 μl and separated using a linear gradient of buffer B (70% acetonitrile in 0.085% TFA) in 0.1% TFA: 5% buffer B (0–5 min), 5% to 55% buffer B (5–50 min), 55% to 75% buffer B (50–55 min) and 75% to 95% buffer B (55–65 min) with a flow-rate of 200 μl/min. Absorption was monitored at 210 nm. The fractions were collected manually to assure accurate peak recoveries and stored at −20°C until analyzed.

2.7. Protein sequencing

Fractions recovered from the HPLC were loaded directly onto one of three automated protein sequencers (Model 470-gas phase 900A data acquisition system, 473 pulse liquid, or 477-pulse liquid, Applied Biosystems). Fractions were applied to glass fiber discs coated with 3.6 mg of polybrene and sequenced using standard Edman N-terminal chemistry. Sequences were obtained by visual inspection of each chromatogram.

3. Results

3.1. Gel elution of protein after SDS–PAGE

Proteins were concentrated either by standard TCA precipitation or by a modified TCA precipitation method for low protein concentrations [12]. The protein mixture was then resolved by SDS–PAGE. The gel was fixed in methanol and acetic acid and stained with Coomassie blue to visualize the protein bands. Fixation and staining of the gel caused passive elution of the proteins from the gel to be inefficient, probably due to the formation of insoluble protein aggregates. This problem was more pronounced for high-molecular-mass proteins.

To improve the yield of protein obtained by passive elution from the gel, we tested the effects of SDS and Lys-C protease. SDS was chosen because it would aid in resolublization of the protein. Lys-C protease was chosen because it would digest the protein into smaller peptides that were easier to elute. Lys-C is also relatively insensitive to SDS, small enough to penetrate the gel matrix [13] and has cleavage sites (Lys-X) that are a subset of the cleavage sites (Lys-X and Arg-X) for trypsin, which could then be used for final digestion of the protein.

After fixation and staining, protein bands were cut out of the gel, washed with water and diced into chips. The gel chips were dehydrated by SpeedVac to facilitate rehydration in buffer A containing different concentrations of SDS with or without Lys-C protease. The eluate was then analyzed by SDS–PAGE in 12.5% polyacrylamide so that either the intact protein or its proteolytic fragments could be resolved. Fig. 1 shows the results with β-galactosidase. The addition of SDS alone to buffer A permitted partial elution of the protein, but only if high concentrations were used. A concentration of 0.01% SDS yielded almost no protein, while 0.1% SDS yielded only about 20% of the starting material (Fig. 1, lanes 1–6). By contrast, the addition of Lys-C protease resulted in a greatly improved yield, even in 0.01% SDS (Fig. 1, lanes 7–9). The yield appears to be about 80–90% when the SDS concentration was 0.03% or greater (Fig. 1, lanes 10–18). The elution yield with 0.04% SDS and Lys-C protease was also measured by amino acid analysis of the eluate and of the gel chips after elution, with 90% of β-galactosidase protein being recovered in the eluate. The elution yield was not significantly affected by gel thickness or polyacrylamide concentration. Gels with thicknesses of 0.75 and 1.5 mm and with 6 and 15% polyacrylamide gave similar recoveries (data not shown).

3.2. Removal of SDS from the protein sample

When 10–100 μg of protein was eluted from gel chips in buffer containing 0.04% SDS (40–100 μg of SDS), the SDS in the eluate was enough to

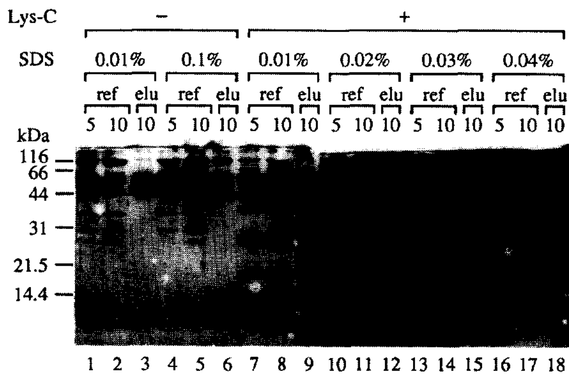


Fig. 1. Improvement of gel elution with SDS and Lys-C protease. To test for the effects of SDS and Lys-C protease on gel elution, β-galactosidase (12 μg, 100 pmol) was resolved by SDS–PAGE, visualized by Coomassie blue staining and cut out in the form of a gel chip. The protein was eluted from the gel in different concentrations of SDS (0.01–0.1%) in the presence or absence of Lys-C protease (0.1 μg). The eluate (elu) was dried down by SpeedVac, resuspended in a volume of 100 μl, of which 10 μl was resolved by SDS–PAGE in a 12.5% polyacrylamide gel and visualized by silver staining (lanes 3, 6, 9, 12, 15 and 18). To estimate recovery from the gel elution, β-galactosidase (12 μg) in a 100-μl volume was incubated with corresponding concentrations of SDS in the presence or absence of Lys-C protease, so that 10 μl and 5 μl were resolved as reference (ref) standards for 100% (lanes 2, 5, 8, 11, 14 and 17) and 50% (lanes 1, 4, 7, 10, 13 and 16) recovery, respectively. The gels was fixed in 50% methanol and 12% TCA overnight and then silver stained as described previously [6]. The recoveries in the eluates were estimated by visual comparison of the densities of the peptide bands between reference and eluate lanes for full length 116 kDa β-galactosidase in the Lys-C (−) lanes and for peptides smaller than 40 kDa in the Lys-C (+) lanes. Each reference lane shows different peptide band patterns because the Lys-C protease digestion was affected by SDS.

partially inhibit trypsin digestion and completely interfere with separation of the tryptic peptides by reversed-phase HPLC (Fig. 3).

Of the many methods described for removing SDS from protein solutions [9,10,14,15], ion-exchange and organic extraction were of particular interest, because both have been shown to remove SDS efficiently enough to allow the separation of peptides by reversed-phase HPLC. We tested both methods for yields with small amounts of protein. Lys-C-digested β-galactosidase (10 μg, 90 pmol) was mixed with 40 μg of SDS and either passed through an AG11A8 ion-exchange column or extracted with a heptane–isoamyl alcohol solution. The samples were then concentrated by SpeedVac and the yield was analyzed by SDS–PAGE (Fig. 2). The ion-exchange column yielded less than 50% recovery, but heptane–isoamyl alcohol gave essentially complete recovery, even after two successive extractions.

Organic extraction of SDS from the protein sample permitted successful digestion with trypsin and resolution by HPLC. When 100 pmol of denatured β-galactosidase were digested with trypsin in 0.04% SDS and loaded onto a reversed-phase HPLC column without heptane–isoamyl alcohol extraction, almost no peaks were observed (Fig. 3A). If the trypsin-digested sample was extracted with heptane–

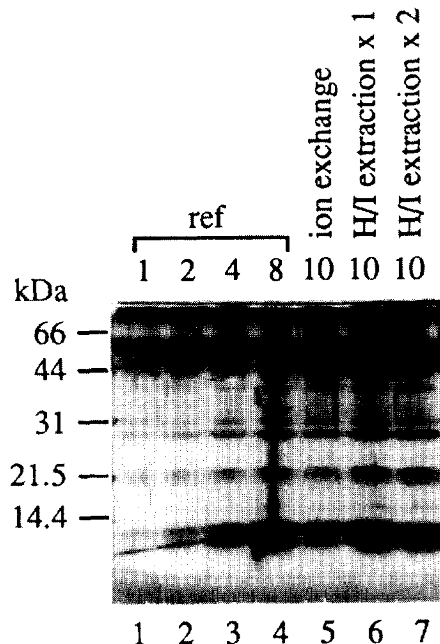
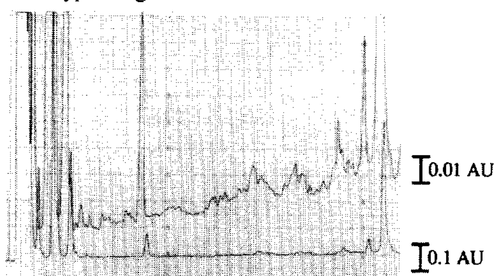
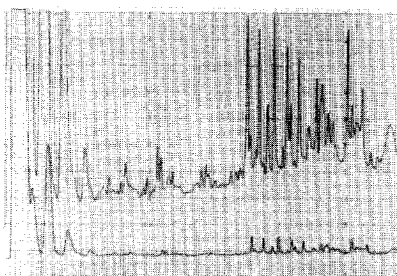


Fig. 2. Comparison of heptane–isoamyl alcohol extraction and ion-exchange chromatography. Each sample containing 10 μg of β-galactosidase, partially digested with Lys-C protease in 0.04% SDS, was either passed over an AG11A8 ion-exchange column (lane 5), or extracted once (lane 6) or twice (lane 7) with heptane–isoamyl alcohol (H/I extraction). The samples were dried down by SpeedVac, resuspended in a volume of 50 μl, of which 10 μl was resolved by 12.5% SDS–PAGE and visualized by silver staining. To estimate recovery, a control sample containing 10 μg of β-galactosidase partially digested with Lys-C protease in 0.04% SDS was dried down and resuspended in 50 μl, so that 1, 2, 4 and 8 μl were resolved as reference (ref) standards for 10, 20, 40 and 80% recovery (lanes 1–4, respectively). The peptide recovery was estimated by visual comparison of the densities of the peptide bands that were smaller than $40 \cdot 10^3$.

A. Trypsin digestion > HPLC



B. Trypsin digestion > H/I extraction > HPLC



C. H/I extraction > trypsin digestion > HPLC

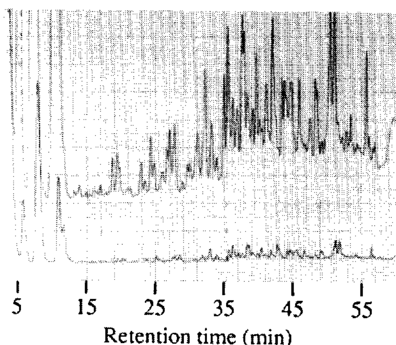


Fig. 3. Effect of heptane–isoamyl alcohol extraction on trypsin digestion and HPLC. Three different protocols were followed to test the efficacy of heptane–isoamyl alcohol extraction for improving trypsin digestion and HPLC of the peptides. In each case, β-galactosidase (100 pmol) was denatured with DTT, urea and iodoacetamide in 0.04% SDS. (A) The denatured protein was digested with trypsin and then resolved by reversed-phase HPLC. (B) The denatured protein was digested with trypsin, extracted to remove SDS and then resolved by HPLC. (C) The denatured protein was extracted, digested with trypsin and then resolved by HPLC. The upper and lower tracings in each panel were plotted at 0.1 and 1.0 absorption units (AU) full scale, respectively.

isoamyl alcohol prior to HPLC, a large number of peaks suitable for sequencing were obtained (Fig. 3B). Thus, the extraction removed sufficient amounts of SDS to restore the resolving power of HPLC. Furthermore, if the protein sample in 0.04% SDS was extracted first, and then digested with trypsin, additional peaks were observed (Fig. 3C). Thus, the heptane–isoamyl alcohol extraction permitted improved trypsin digestion. In fact, the HPLC profile was virtually identical to that obtained from trypsin digestion of β-galactosidase in the absence of SDS (compare Fig. 3C and Fig. 4C).

3.3. Trypsin digestion of the gel purified protein and HPLC of the peptides

Protein eluted from gel chips by digestion with Lys-C protease in 0.04% SDS was only partially digested by Lys-C, which was not adequate enough to give peptide peaks on HPLC suitable for sequencing. Therefore, the protein fragments eluted from the gel were extracted with heptane–isoamyl alcohol to remove the SDS, denatured with urea, reduced with DTT, alkylated with iodoacetamide and digested with trypsin. The sample was then loaded onto a reversed-phase HPLC column to resolve the tryptic peptides.

A polyacrylamide gel chip containing no protein produced no significant background on HPLC (Fig. 4A). A polyacrylamide gel chip containing 100 pmoles of β-galactosidase produced a large number of sharp peaks (Fig. 4B). In fact, the profile was very similar to the HPLC profile from β-galactosidase digested in solution (Fig. 4C). The average yield was about 60%, with some peptide peaks showing significantly more recovery and others showing less recovery. Apparently, there was heterogeneous loss of peptides during elution and extraction, perhaps from incomplete digestion in situ in the gel with Lys-C or the loss of hydrophobic peptides into the organic phase during the heptane–isoamyl alcohol extraction.

3.4. Application to other proteins

The EDS method was tested for β-galactosidase because its molecular mass of $116 \cdot 10^3$ was relatively large and therefore would be a stringent test for

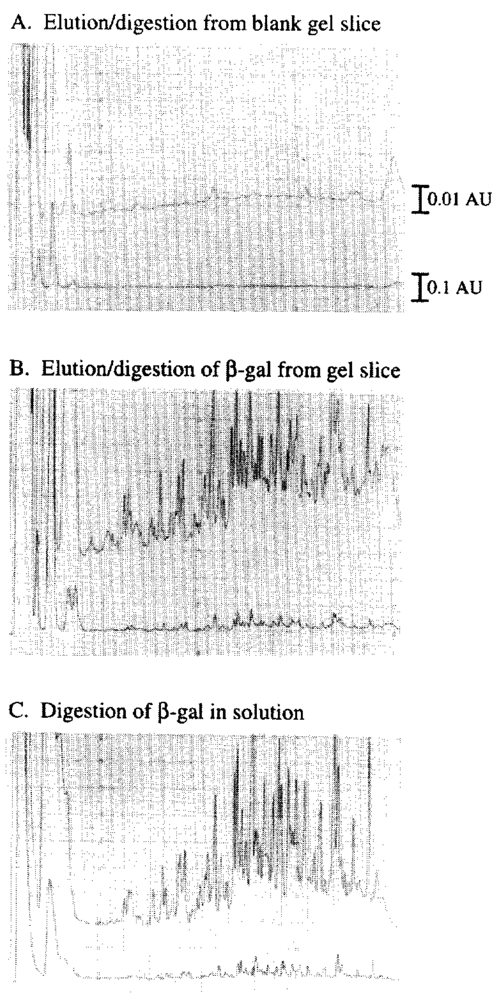


Fig. 4. HPLC analysis of protein after elution–digestion from a gel. To test the efficiency of the EDS method, 100 pmol of β-galactosidase (β-gal) was resolved by SDS–PAGE and subjected to elution–digestion. (A) HPLC profile of a blank gel chip subjected to the EDS method, showing the background signal. (B) HPLC profile of β-galactosidase subjected to the EDS method. (C) HPLC profile of β-galactosidase digested in solution with trypsin. The upper and lower tracings were plotted at 0.1 and 1.0 absorption units (AU) full scale, respectively.

successful elution from the gel. To test the general applicability of the method, we applied it to a number of additional proteins with a wide range of molecular masses; myosin ($200 \cdot 10^3$), XPE-BF ($125 \cdot 10^3$), phosphorylase *b* ($94 \cdot 10^3$), p93 ($93 \cdot 10^3$), BSA ($66 \cdot 10^3$) and ovalbumin ($45 \cdot 10^3$). All proteins subjected to the EDS method yielded sharp peptide peaks on HPLC, with average peak heights proportional to the amount of protein analyzed (Fig. 5A–F). Although myosin had a very high molecular mass, it was successfully eluted from the gel. Many of the peaks for myosin were clearly overlapping because of the large number of tryptic peptides, but there were a significant number of isolated peaks, which could yield useful sequence information. To avoid this problem for very high molecular mass proteins, Lys-C protease rather than trypsin may be used in the second digestion.

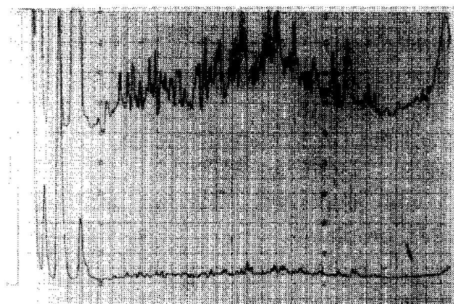
The EDS method was tested for the amount of amino acid sequence information that could be obtained from small amounts of protein. Thus, we analyzed 25 pmol of β-galactosidase, 20 pmol of XPE-BF and 10 pmol of a 93 kDa protein (p93) that copurified with XPE-BF [6,16]. Eleven peptide peaks were chosen for analysis and all eleven yielded sequence information (Fig. 5). The percentage yields from peptides that could be sequenced were between 25% and 80%, with an average yield of 55% (Table 2). This is consistent with an average yield of 60%, estimated by comparing HPLC peak heights to those obtained for 100 pmol of β-galactosidase (Fig. 4).

4. Discussion

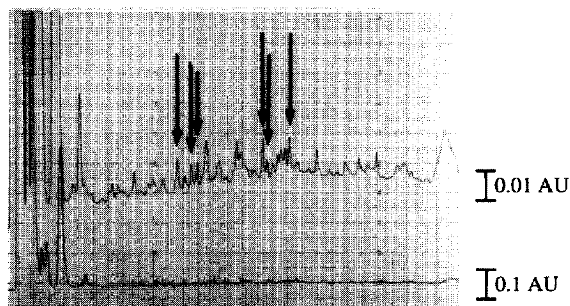
The elution–digestion–sequencing (EDS) method was successful in reliably obtaining amino acid sequence information from gel-purified protein. Efficient elution was achieved by digestion with Lys-C protease and elution with SDS. The efficiency of elution was independent of molecular mass, since Lys-C digested high-molecular-mass proteins into fragments that could be eluted. SDS and other contaminants in the eluate were successfully removed by heptane–isoamyl alcohol extraction, allowing the eluted protein to be re-digested in solution with trypsin. The overall yield was approximately 60%.

For high-molecular-mass proteins, trypsin digestion may yield such a large number of peptides that many of the peaks on HPLC might contain more than one peptide. This was seen for the case of the 125 kDa protein, XPE-BF (Table 2), in which four of the six HPLC peaks contained two or more peptides. This problem may be circumvented by

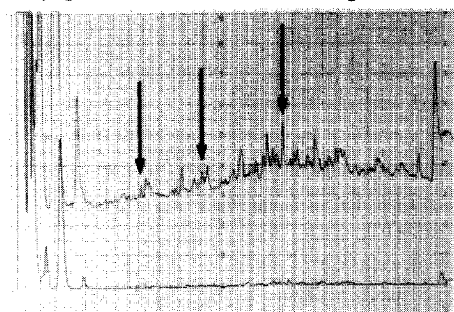
A. Myosin (200 kDa, 60 pmol)



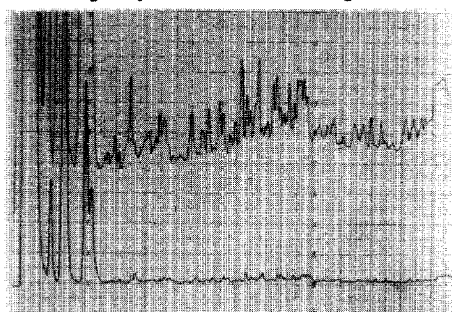
B. XPE-BF (125 kDa, 20 pmol)



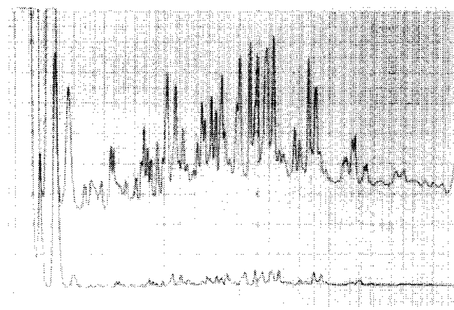
C. β-galactosidase (116 kDa, 25 pmol)



D. Phophorylase b (94 kDa, 53 pmol)



E. BSA (66 kDa, 125 pmol)



F. Ovalbumin (45 kDa, 290 pmol)

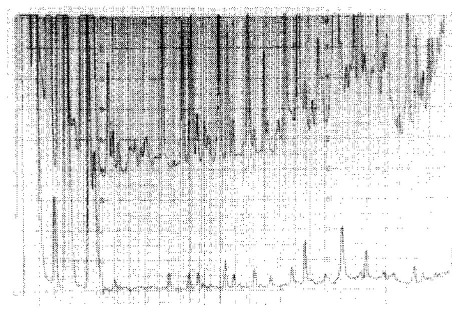


Fig. 5. Elution–digestion of different proteins. Different proteins representing a range of molecular masses and in a range of different amounts were subjected to the EDS method and resolved by HPLC. The arrows represent peptide peaks that were subjected to microsequencing analysis, the results of which are shown in Table 2. The upper and lower tracings were plotted at 0.1 and 1.0 absorption units (AU) full scale, respectively.

eliminating the trypsin digestion in favor of a second Lys-C digestion in a solution with Lys-C at a ratio of 1:10 (Lys-C/protein, w/w).

The EDS method obtained sequences for proteins with a molecular mass as large as $200 \cdot 10^3$, which in the past have proven difficult to sequence, and from as little as 10 pmol of protein loaded onto the gel. The method required only the standard equipment for amino acid sequencing. Thus, most of the procedure may be performed by the investigators who purify the protein (steps 1–11 in Table 1). The protease-digested protein can then be analyzed in a central

Table 2
Sequencing results for proteins analyzed by the EDS method

Protein (molecular mass)	Starting material (pmol)	Cycle 1 yield[a] (pmol)	Yield[b] (%)	Amino acid sequence[c]	Sequence position
β-galactosidase ($116 \cdot 10^3$)	25	ND	ND	xSEEAR	38
		12	48	xVQPNATA	700
XPE-BF ($125 \cdot 10^3$)	20	15	75	TYEVSLR	191
		10	50	YLAIAPVxxDF	244
		ND	ND	xxMLLLE	279
		12	60	ALYYLQIHPQELR	514
		ND	ND	xxNTYFIVGPAMVYPE	823[d]
		11	55	DPNTYFIVGTAMVYPEEAEP	823[d]
		7	35	xVEELTRI	1131
		5	25	xAVYSMVEF	867
		7	35	MQEVVANLQYDDG	1104
		16	80	REATADD	1121
p93 ($93 \cdot 10^3$)	10	6	60	IVVFQYS	847
		5	50	xEPATGFIDG	1081

[a] Estimated yield calculated from cycle 1 or by extrapolation of yields from subsequent cycles to cycle 1.
[b] Percentage yield of the starting material available for sequence analysis.
[c] The letter x in the amino acid sequences denotes an amino acid that could not be identified.
[d] The peptide starting at position 823 in XPE-BF was found in two different HPLC peaks and contained either threonine or proline at position 832, possibly representing a polymorphism in XPE-BF, since the protein was purified from ten pooled placentas.
ND: Not determined (because the yields from the first two cycles of sequencing were lower than the yields from subsequent cycles, making an estimate of the yield from cycle 1 unreliable).

facility, in which the peptides are separated by HPLC and sequenced by Edman degradation.

There are several reports that resolve proteins by SDS–PAGE and digest proteins in situ in the gel [13,17,18]. Ward et al. [17] identified proteins by staining with Coomassie blue, digested the proteins in situ in the gel with trypsin and then extracted the peptides from the gel with TFA. They tested small proteins, β-lactoglobulin and interleukin-6 (with molecular masses of $18 \cdot 10^3$ and $22 \cdot 10^3$, respectively) in large amounts (1100 and 920 pmol). They found that the yield using this method was superior to that obtained from the digestion of proteins electroblotted onto PVDF membrane. However, despite the favorable circumstance of abundant low-molecular-mass proteins, their yield averaged only 20% (ranging from 3–67%), significantly worse than for the EDS method.

Rosenfeld et al. [18] reported improved results by modifying the standard Coomassie staining procedure. The concentrations of methanol and acetic acid were reduced to minimize fixation of the proteins into the gel matrix. The gel was washed in 50% acetonitrile to remove contaminants that interfere with enzymatic digestion, was partially dried with significant shrinkage and then was rehydrated in a trypsin solution. The average yield was about 65% when applied to proteins with a molecular mass smaller than $50 \cdot 10^3$ in amounts of about 100 pmol, but no data was shown for larger proteins in smaller amounts. The authors noted that complete dehydration of the gel led to severely reduced yields. However, Hellman et al. [19] reported better results when the gel was completely dehydrated, but did not explain the discrepancy. We tested the method of Rosenfeld et al. [18] with a series of proteins ranging in molecular mass from 14 to $200 \cdot 10^3$ and found several problems (data not shown). First, the method was sensitive to the ratio of trypsin to substrate. Although a ratio of 1:4 (w/w) was sufficient for digestion of many proteins, a more conventional ratio of 1:10 (w/w) was not adequate for efficient digestion. Second, the method failed to work well for some proteins, failing for RNAse A ($14 \cdot 10^3$) and phosphorylase *b* ($94 \cdot 10^3$), while succeeding for ovalbumin ($45 \cdot 10^3$), BSA ($66 \cdot 10^3$), β-galactosidase

($116 \cdot 10^3$) and myosin ($200 \cdot 10^3$). Third, the method was sensitive to high polyacrylamide concentrations. Proteins larger than $60 \cdot 10^3$ could be digested in 7.5%, but not in 12.5%, polyacrylamide gels. Thus, this method gives variable results.

Kawasaki et al. [13] used a standard Coomassie blue staining procedure and digested the proteins in situ in the gel with Lys-C protease, thus avoiding the problem of poor trypsin digestion, since Lys-C is less sensitive to SDS. To facilitate elution from the gel, the Lys-C digestion was performed in the presence of 0.1% SDS. The SDS was then removed by a DEAE ion-exchange column. Recovery was excellent in the test case of the relatively large protein, phosphorylase *b* ($94 \cdot 10^3$). However, the amount of protein was relatively high (110 pmol), and several peptide peaks were lost due to SDS contamination. In order to overcome the inhibitory effect of 0.1% SDS, the method requires that the Lys-C concentration be kept greater than 2 μg/ml. In fact, the digestion of apomyoglobin succeeded only when the Lys-C concentration was increased to 10 μg/ml [20]. As a result, this method becomes much less efficient when smaller amounts of protein are available for analysis, if the enzyme–substrate ratio is to be kept less than 1:10. Furthermore, digestion with Lys-C produces only half the number of peptides as trypsin, limiting the amount of amino acid sequence information that can be obtained.

In conclusion, the previous in situ gel digestion methods have been inconsistent in generating peptide sequences from proteins with a molecular mass larger than $100 \cdot 10^3$ or in amounts as small as 10 pmol. The EDS method succeeds because elution from the gel and digestion of the protein are both highly efficient. This is achieved by utilizing two separate proteolysis steps: one digestion in situ in the gel to facilitate elution from the gel; and a second digestion in solution to facilitate efficient cleavage. Therefore, the EDS method takes longer than the other methods described above. On the other hand, we have now applied the EDS method to obtain peptide sequence information from each of the 25 proteins analyzed so far (data not shown). The method appears to be robust enough to be generally applicable to proteins of both low and high molecular masses in amounts as small as 10 pmol. Thus, the EDS method will be particularly valuable in cases where purification of the targeted protein is difficult and reliability of the sequencing method is of paramount importance.

Acknowledgments

This work was supported by NIH grant CA 44949 and by Graham and Jane Nissen. The authors thank Jim Kenny and Young Moo Lee for valuable discussions and technical expertise, Kim Rathmell, Vaughn Smider, Joseph Liao and Brian Feldman for critically reading the manuscript.

References

[1] J. Brown and W. Roberts, J. Biol. Chem., 251 (1976) 1009.
[2] R. Aebersold, J. Leavitt, R. Saavedra, L. Hood and S. Kent, Proc. Natl. Acad. Sci. U.S.A., 84 (1987) 6970.
[3] G. Bauw, J. van Damme, M. Puype, J. Vandekerckhove, B. Gesser, G. Ratz, J. Lauridsen and J. Celis, Proc. Natl. Acad. Sci. U.S.A., 86 (1989) 7701.
[4] J. Fernandez, M. DeMott, D. Atherton and S. Mische, Anal. Biochem., 201 (1992) 255.
[5] R. Aebersold, in P. Matsudaira (Editor), A Practical Guide to Protein and Peptide Purification for Microsequencing, Academic Press, New York, 1989, p. 71.
[6] B.J. Hwang and G. Chu, Biochemistry, 32 (1993) 1657.
[7] U. Laemmli, Nature, 227 (1970) 680.
[8] S. Moore and W. Stein, in S. Colowick and N. Kaplan (Editors), Methods in Enzymology, Academic Press, New York, 6 (1963) 819.
[9] A. Bosserhoff, J. Wallach and R. Frank, J. Chromatogr., 473 (1989) 71.
[10] O. Kapp and S. Vinogradov, Anal. Biochem., 91 (1978) 230.
[11] K. Stone, M. LoPresti, J. Crawford, R. DeAngelis and K. Williams, in P. Matsudaira (Editor), A Practical Guide to Protein and Peptide Purification for Microsequencing, Academic Press, New York, 1989, p. 31.
[12] B.J. Hwang and G. Chu, Biotechniques, 20 (1996) 1982.
[13] H. Kawasaki, Y. Emori and K. Suzuki, Anal. Biochem., 191 (1990) 332.
[14] W. Konigsberg and L. Henderson, in C. Hirs and S. Timasheff (Editors), Methods in Enzymology, Academic Press, New York, 91 (1983) 254.
[15] D. Wessel and U. Flugge, Anal. Biochem., 138 (1984) 141.
[16] B.J. Hwang, J. Liao and G. Chu, Mutation Res., 362 (1996) 105.

[17] L.D. Ward, G.E. Reid, R.L. Moritz and R.J. Simpson, J. Chromatogr., 519 (1990) 199.
[18] J. Rosenfeld, J. Capdevielle, J. Guillemot and P. Ferrara, Anal. Biochem., 203 (1992) 173.
[19] U. Hellman, C. Wernstedt, J. Gonez and C. Heldin, Anal. Biochem., 224 (1995) 451.
[20] H. Kawasaki and K. Suzuki, Anal. Biochem., 186 (1990) 264.