ELSEVIER

Journal of Chromatography A, 736 (1996) 291–302

JOURNAL OF CHROMATOGRAPHY A

Large-scale amino-acid analysis for proteome studies

Jun X. Yan[a], Marc R. Wilkins[a], Keli Ou[a], Andrew A. Gooley[a], Keith L. Williams[a,*], Jean-Charles Sanchez[b], Olivier Golaz[b], Christian Pasquali[b], Denis F. Hochstrasser[b]

[a]*Macquarie University Centre for Analytical Biotechnology, Macquarie University, Sydney, NSW 2109, Australia*

[b]*Central Clinical Chemical Laboratory and Medical Computing Centre, University of Geneva, CH-1121 Geneva 14, Switzerland*

Received 11 July 1995; revised 14 November 1995; accepted 20 November 1995

Abstract

Amino-acid analysis is a relatively new method for identification of proteins separated by two-dimensional gel electrophoresis and blotted onto polyvinylidene difluoride (PVDF) membranes. This article describes modified amino-acid analysis methods for this purpose. Streamlined sample handling is a key feature of the process. To minimise sample manipulation, a single vial is used for hydrolysis and the protein hydrolysate on PVDF membrane is extracted by a one-step procedure. The hydrolysate should not be stored for long periods before analysis. Applications of the technique are presented to demonstrate the identification procedure. This approach is the most cost-effective and time-effective first step in mass protein screening for a large-scale proteome project.

Keywords: Proteomes; Amino acids; Proteins

1. Introduction

Proteomes are the entire PROTEin complement expressed by the genOME of a simple organism, or the protein complement of a tissue or cell [1,2]. The proteins are separated by two-dimensional (2-D) gel electrophoresis and identified by a variety of techniques including Edman degradation and antibody cross reaction [3–7]. However, identification of protein spots is the rate-limiting step and this has led to renewed interest in amino-acid analysis as a means of rapid protein identification [1,2,8–13]. Amino-acid analysis represents a useful first step in protein identification as it is cost-effective and time-effective in large-scale studies on proteomes [1]. For well-studied organisms (e.g., *Escherichia coli*) for which a large segment of the genome has been sequenced, 70% of proteins are confidently and correctly identified by a combination of amino-acid analysis and p*I* and apparent mass (estimated from the 2-D gel) [1]. Identification of the remaining proteins can be achieved by a hierarchical approach [2] using, for example, peptide-mass fingerprinting; a combination of amino-acid analysis and peptide-mass fingerprinting; N-terminal sequence analysis; internal protein microsequencing; microsequencing by mass spectrometry (electrospray ionisation and post-source decay MALDI–TOF).

For successful identification using the hierarchical approach, an accurate amino-acid composition for each protein spot is critical. This accuracy can only be obtained on the basis of high resolution of 2-D PAGE (polyacrylamide gel electrophoresis) separation, a high efficiency of electroblotting, and high-

*Corresponding author.

0021-9673/96/$15.00
SSDI 0021-9673(95)01285-0

quality amino-acid analysis. There have been dramatic advances in 2-D electrophoresis with the immobilised pH gradient technique providing highly reproducible 2-D maps [14,15]. By comparison, amino-acid analysis of proteins from PVDF membranes for protein identification purposes is poorly understood. It is reported to be more difficult to conduct amino-acid analysis from PVDF-bound protein spots than from pure proteins in solution [16–21]. This perhaps explains why the technique is not as popular as may be expected. Where possible we believe that it is preferable to develop techniques for accurate amino-acid analysis, rather than to rely on error weighting for obtaining the amino-acid composition, although some data manipulation may be needed [10–12]. Furthermore, if large-scale proteome studies are to be conducted, minimising the analysis time will allow high sample throughput. To date, there has been little or no attention paid to the optimisation of conditions for doing amino-acid analysis from PVDF-bound proteins in order to identify them. One important feature of protein spots separated by 2-D electrophoresis is their high purity, free of many of the components that have contaminated the analysis of protein bands separated by 1-D electrophoresis.

We report here a study of protein hydrolysis, and amino-acid analysis of proteins from 1-D or 2-D PAGE with Fmoc pre-column derivatisation in an automated analysis system. Data are presented to demonstrate amino-acid analyses of high quality, which result in protein identification.

2. Experimental

2.1. Materials

9-Fluorenylmethyl chloroformate, hydroxylamine hydrochloride, 2-(methylthio)-ethanol, 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), bovine serum albumin (BSA), egg albumin (Ovalbumin), trypsinogen and β-lactoglobulin were obtained from Sigma. Amino-acid calibration standard A "Sepramar", ammonium monohydrogen phosphate, anhydrous ammonium dihydrogen phosphate, boric acid, sodium hydroxide, phenol (analytical-reagent grade), 11.4 *M* hydrochloric acid (Aristar grade) and sodium dodecyl sulphate (SDS) were purchased from BDH. Acetic acid (analytical-reagent grade) was obtained from Mallinckrodt. Acetonitrile (HPLC grade), methanol (HPLC grade) and ethanol (HPLC grade) were purchased from Ajax (Australia). Triethylamine (TEA) was obtained from Pierce. Trifluoroacetic acid (TFA) was purchased from LKB Biochem. Acrylamide, N,N′-methylenebisacrylamide, Immobilon-P transfer PVDF membranes (0.45 μm) and a Milli-Q water system were obtained from Millipore. Trans-Blot transfer PVDF membranes (0.2 μm) were obtained from Bio-Rad.

2.2. Electrophoresis and electroblotting

Isoelectric focusing procedure: IPG (immobilised pH gradient) 18-cm strips (Pharmacia) were used for the first-dimension electrophoresis following the methods described by Bjellqvist et al. [14,15]. SDS–polyacrylamide slab gels (20×20 cm) (Bio-Rad) were used in the second dimension, and mini-protean II (Bio-Rad) was used for 1-D electrophoresis [22]. Proteins were electrotransferred to PVDF using 10 m*M* CAPS buffer (pH 11.0) in 10% (v/v) methanol [23] and PVDF was stained with 0.1% (w/v) amido black in Milli-Q water and destained in Milli-Q water (modified from Sanchez et al. [24]). The PVDF membrane was then dried under nitrogen to improve contrast of protein spots. Protein spots were cut out and stored at −20°C until use.

2.3. Hydrolysis methods

2.3.1. Hydrolysis procedure

Dissolved protein samples (50–150 pmol) were added into 0.7-ml glass autosampler vials and dried under vacuum prior to hydrolysis. PVDF spots carrying protein were put into vials with a clean needle.

The hydrolysis vessel has been specifically designed for hydrolysis of small amounts of protein [25]. The vessel incorporates a belly-bottomed flask, which is threaded at the top to match a Corning SVL30 fitting. The SVL30 fitting has a permanently attached vacuum tap (Fischer and Porter) with a teflon piston that incorporates a replaceable rubber O-ring. A Kal-rez O-ring (part 221010 from Applied

Biosystems, Division of Perkin-Elmer) is used to seal the belly bottom to the top, inside the SVL30 fitting.

Into the bottom of the hydrolysis vessel was placed 400 μl of 5.7 *M* or 11.4 *M* HCl, a crystal of phenol (about 0.1%, w/v) and up to 15 autosampler vials. The vessel was assembled tightly, alternately evacuated and flushed with argon twice, and finally evacuated for 10 s and sealed. Evacuated vessels are then housed in a solid aluminium cylinder of 8 cm radius and 13 cm height, which has been drilled out to 4.6 cm radius and 9.5 cm height to hold the hydrolysis vessel, with the hollow just deep enough so that the base of the hydrolysis vessel can rest on the metal inside the holder. These metal holders are kept in a 155°C oven to maintain the correct temperature for hydrolysis. Hydrolysis was for 1 h at 155±5°C. Following hydrolysis, vessels were opened inside a fume hood, and vials were moved to a vacuum drier for 10 min to remove residual acid vapour.

2.3.2. Extraction of amino acids from PVDF

A one- or two-step extraction procedure was used. The one-step method is as follows: 170 μl extraction solution (50 μl Milli-Q water, 100 μl acetonitrile and 20 μl 0.1% TFA (v/v)) was added into vials containing a PVDF spot, sealed with parafilm and sonicated for 10 min. The PVDF was removed with a clean needle and the solution was lyophilised. The two-step method [17] is as for the one-step method with the addition of a further neutralisation following extraction. The neutralisation involves resuspending the amino acids in 20 μl of 100% TEA–ethanol–water (2:2:1, v/v/v; pH 10.0) and evaporating to dryness. Samples were prepared for amino-acid analysis by adding 10 μl of 250 m*M* sodium borate buffer (pH 8.8).

2.4. Amino-acid analysis

Amino-acid analysis using pre-column Fmoc derivatisation was undertaken as described in an accompanying paper [26]. The most attractive feature of Fmoc amino-acid analysis is that the derivatisation chemistry is fully automated. The literature has neglected the utility of Fmoc amino-acid analysis despite the developments involving mono-derivatisation of histidine and tyrosine and the elimination of an extraction step [27,28].

The amount of each amino acid present (pmol) in a protein hydrolysate was calculated from a known standard for each amino acid, where the pmol of sample amino acid=(sample peak area/standard peak area)×(pmol of standard). The accuracy for each amino acid was validated by calculating % error, and the overall accuracy of the composition as average % error, where: % error=100×(experimental value−known value)/(known value) and average % error=(Σ/% error for 16 amino acids/)/16 [16,18].

2.5. Protein identification by database searching

Following amino-acid analysis of an unknown protein spot separated by 1-D or 2-D gel, its amino-acid composition, p*I* and M_r (where applicable) were matched against entries in the SWISS-PROT database according to Wilkins et al. [1] through the ExPASy computer program (AAcompID) via the World Wide Web network (http://expasy.hcuge.ch/ch2d/aacomp2.html). Constellation 2 was used (Asx, Glx, Ser, His, Gly, Thr, Ala, Pro, Tyr, Arg, Val, Met, Leu, Ile, Phe and Lys are considered, Asx=Asp+Asn, Glx=Gln+Glu, Cys and Trp are not considered), either with or without p*I* and M_r windows of ±0.5/±1 units and ±20%, respectively. The unknown protein can be identified either by single-species matching if the species is known, or by cross-species matching for all species. To demonstrate the actual quality of our amino-acid analysis, no calibration or normalisation of data was used before matching in this study. However, we routinely calibrate with a known protein in our protein identification studies [1].

3. Results and discussion

3.1. Protein hydrolysis of PVDF-bound spots at pmol levels

3.1.1. Indication of a successful hydrolysis

When doing protein hydrolysis, it is of benefit to have some indication if the hydrolysis has been successful. A good seal is obvious by the sudden release of vapour once the vessel's seal is undone. In

addition, we have noted that PVDF spots stained with amido black, which appear blue/black before hydrolysis, and appear completely white after a successful hydrolysis. This is a useful indication that hydrolysis has been successful. This loss of colour is not observed when the PVDF blot is stained with either Coomassie blue or Ponceau S (data not shown).

3.1.2. Acid concentration for hydrolysis

The concentration of hydrochloric acid is recommended to be 5.7 *M* for 22–24 h hydrolysis at 110°C or 5.7 *M* for 1–2 h at 155°C with non-PVDF-bound proteins, although it has been suggested that the shortened hydrolysis time may cause the incomplete release of amino acids from peptide linkage of valine and isoleucine [29]. Another study used a 1:2 mixture of trifluoroacetic and 11.4 *M* HCl at 155°C for 1 h [30]. It was not known what the higher concentration of acid would do to hydrolysis of PVDF-bound protein. Accordingly, we investigated the use of 11.4 *M* hydrochloric acid for hydrolysis at 155°C for 1 h. All protein hydrolysates were extracted by the one-step method. We found that hydrolysis with 11.4 *M* HCl provided good results for PVDF-bound proteins, as well as for non-PVDF-bound proteins and we did not encounter difficulties with valine and isoleucine. We speculate that proteins separated by 2-D PAGE are fully hydrolysed due to their open conformation on PVDF. The same results were also obtained using 5.7 *M* HCl hydrolysis. The amino-acid compositions of non-PVDF-bound proteins and PVDF-bound proteins hydrolysed by 5.7 *M* or 11.4 *M* HCl are shown in Tables 1 and 2, respectively. These proteins were BSA (average of 15 pmol hydrolysed), trypsinogen (average 70 pmol hydrolysed) and ovalbumin (average 20 pmol hydrolysed). The mean average % errors obtained for trypsinogen, BSA and ovalbumin are shown in Table 3. The accuracy of analysing the amino-acid composition of PVDF-bound proteins is as good as or better than other studies where the mean average % error from 30 sites was found to be 27.8 and protein used was 1 μg (59 pmol) of horse apomyoglobin absorbed onto PVDF membrane [18], where no electrophoresis was involved. Hence, no Gly contamination and no Met oxidation were likely to occur [31]. In addition, to minimise the analysis cost, and the possibility of introducing contamination, we do not dilute 11.4 *M* HCl. We conclude that 11.4 *M* HCl is suitable for acid hydrolysis (1 h, 155°C) of PVDF-bound proteins for the purpose of protein identification by amino-acid composition. In a mass screening study, cost-effectiveness and time-effectiveness become an important issue. Therefore, using 11.4 *M* HCl for hydrolysis is a reasonable option.

3.2. Modification of extraction method from PVDF

If large numbers of samples are to be analysed, the key issue for sample preparation is to minimise sample handling. Each time a sample is manipulated it is subject to losses or contamination; extra time is also introduced into the procedure. The extraction procedure outlined in Section 2.3, using 0.1% (v/v) TFA–acetonitrile–water (2:10:5, v/v) and 100% TEA–ethanol–water (2:2:1, v/v; pH 10.0), is referred as the "two-step method". The modification of this procedure by using only 0.1% v/v TFA–acetonitrile–water (2:10:5, v/v) for extraction is referred to as the "one-step method". These two methods were compared to investigate the recovery of protein from SDS-PAGE via PVDF to amino-acid analysis, and to assess the accuracy of the analysis. β-Lactoglobulin was used for the investigation with a starting amount of 165 pmol (3 μg) of total protein (calibrated by amino-acid analysis on an equal amount of protein in solution) loaded onto a 1-D SDS-gel.

3.2.1. Recoveries

The amounts of β-lactoglobulin on PVDF hydrolysed were found to be 87.1 pmol for the one-step method and 86.7 pmol for the two-step method. The recoveries from SDS-PAGE to electroblotting were 52.8 and 52.5%, respectively. This indicated that the two methods achieved similar recovery of amino acids. These results compare favourably with other laboratories' results, where recoveries were reported from 28 to 53% for 2 μg or more of electroblotted proteins hydrolysed and extracted [17,19,20].

3.2.2. Amino-acid composition of PVDF-bound β-lactoglobulin

The amino-acid composition of PVDF-bound β-lactoglobulin was compared with the expected composition calculated from the SWISS-PROT database. Very close matches were shown to the expected

Table 1
Hydrolyses of non-PVDF-bound proteins; 11.4 *M* and 5.7 *M* hydrochloric acid at 155°C for 1 h

Amino acid	Ovalbumin			Trypsinogen			BSA		
	Expected (%)	5.7 *M* HCl $n=7$	11.4 *M* HCl $n=7$	Expected (%)	5.7 *M* HCl $n=6$	11.4 *M* HCl $n=7$	Expected (%)	5.7 *M* HCl $n=7$	11.4 *M* HCl $n=7$
Asx	8.2	9.7	9.1	12.2	12.2	11.4	10.1	10.7	10.5
Glx	12.8	12.7	12.1	6.6	7.4	7.3	14.5	13.8	13.7
Ser	10.1	9.1	8.9	16.0	13.6	13.1	5.1	3.6	3.8
His	1.9	1.8	2.0	1.4	1.6	1.6	2.9	3.4	3.2
Gly	5.1	5.8	5.5	11.7	13.0	11.2	2.9	3.3	3.5
Thr	4.0	4.7	4.7	4.7	5.0	5.2	6.1	4.4	4.8
Ala	9.3	8.8	8.7	6.6	7.4	7.4	8.6	6.8	7.0
Pro	3.7	4.5	4.6	3.8	5.2	4.9	5.1	6.0	6.0
Tyr	2.7	2.8	2.2	4.7	4.5	4.7	3.7	4.5	3.9
Arg	4.0	4.3	3.6	0.9	1.2	1.4	4.2	5.1	4.6
Val	8.2	7.2	7.6	8.5	6.5	7.5	6.6	7.2	7.7
Met	4.3	4.4	4.6	0.9	1.2	1.3	0.7	1.2	1.1
Ile	6.6	5.7	5.9	7.0	6.3	7.0	2.6	2.8	2.9
Leu	8.5	8.7	8.3	6.6	7.4	7.4	11.2	11.8	11.3
Phe	5.3	4.9	6.6	1.4	1.4	1.7	5.0	4.0	3.7
Lys	5.3	4.9	5.8	7.0	6.1	7.0	10.8	11.3	12.3

Table 2
Hydrolyses of PVDF-bound proteins; 11.4 *M* and 5.7 *M* hydrochloric acid at 155°C for 1 h

Amino acid	Ovalbumin			Trypsinogen			BSA		
	Expected (%)	5.7 *M* HCl $n=9$	11.4 *M* HCl $n=9$	Expected (%)	5.7 *M* HCl $n=9$	11.4 *M* HCl $n=7$	Expected (%)	5.7 *M* HCl $n=9$	11.4 *M* HCl $n=7$
Asx	8.2	9.7	10.3	12.2	11.1	10.5	10.1	10.4	10.7
Glx	12.8	13.4	13.5	6.6	7.0	6.6	14.5	13.0	13.8
Ser	10.1	10.6	10.1	16.0	13.7	14.2	5.1	5.5	6.0
His	1.9	1.6	1.5	1.4	1.7	1.6	2.9	2.3	2.2
Gly	5.1	6.8	7.2	11.7	12.8	12.5	2.9	5.1	6.0
Thr	4.0	3.8	4.4	4.7	4.4	4.6	6.1	5.2	5.9
Ala	9.3	9.4	9.3	6.6	7.1	7.2	8.6	8.5	8.8
Pro	3.7	4.6	4.2	3.8	4.8	4.1	5.1	5.9	6.1
Tyr	2.7	2.4	2.2	4.7	4.0	3.8	3.7	3.2	2.4
Arg	4.0	4.0	4.0	0.9	1.0	1.0	4.2	4.8	4.5
Val	8.2	7.0	7.1	8.5	6.6	6.3	6.6	5.2	5.8
Met	4.3	1.7	1.8	0.9	1.1	0.8	0.7	0.7	0.8
Ile	6.6	6.1	5.9	7.0	6.6	6.1	2.6	2.2	2.3
Leu	8.5	8.8	8.5	6.6	7.3	7.0	11.2	9.6	9.8
Phe	5.3	4.7	4.7	1.4	4.1	5.6	5.0	8.7	6.0
Lys	5.3	5.2	5.2	7.0	6.9	7.9	10.8	9.7	8.8

Table 3
Mean average % error for amino-acid analyses and score from AAcompID

Experiment	Protein		Mean average % error	Matching score[a]
(A) Acid concentration test on PVDF-bound proteins	Ovalbumin	11.4 *M* HCl	12.0 (n=9)	8
	Ovalbumin	5.7 *M* HCl	16.3 (n=9)	16
	BSA	11.4 *M* HCl	20.0 (n=7)	22
	BSA	5.7 *M* HCl	23.5 (n=9)	29
	Trypsinogen	11.4 *M* HCl	29.4 (n=7)	30
	Trypsinogen	5.7 *M* HCl	25.6 (n=9)	21
(B) Acid concentration test on non-PVDF-bound proteins	Ovalbumin	11.4 *M* HCl	11.1 (n=7)	18
	Ovalbumin	5.7 *M* HCl	9.2 (n=7)	8
	BSA	11.4 *M* HCl	16.8 (n=7)	18
	BSA	5.7 *M* HCl	15.3 (n=7)	14
	Trypsinogen	11.4 *M* HCl	15.6 (n=7)	16
	Trypsinogen	5.7 *M* HCl	18.4 (n=6)	19
(C) Extraction				
One-step	β-lactoglobulin	PVDF-bound	12.7 (n=4)	8
Two-step	β-lactoglobulin	PVDF-bound	19.6 (n=4)	18
(D) Storage				
Day 0	β-lactoglobulin	non-PVDF-bound	15.5 (n=1)	22
Day 1	β-lactoglobulin	non-PVDF-bound	8.3 (n=1)	5
Day 3	β-lactoglobulin	non-PVDF-bound	8.6 (n=1)	5
Day 7	β-lactoglobulin	non-PVDF-bound	8.9 (n=1)	5
Day 14	β-lactoglobulin	non-PVDF-bound	5.7 (n=1)	4
Day 21	β-lactoglobulin	non-PVDF-bound	6.2 (n=1)	5

(A) for three PVDF-bound standard proteins, using 11.4 or 5.7 *M* HCl during hydrolysis; (B) for three non-PVDF-bound standard proteins, using 11.4 or 5.7 *M* HCl during hydrolysis; (C) for PVDF-bound β-lactoglobulin hydrolysed and subjected to one- or two-step extraction; (D) effect of storage time on β-lactoglobulin.

[a] In all cases the matching to the database led to the protein being identified as the best match.

composition using both of the extraction methods. The average % errors were 12.7 and 19.6 for one-step and two-step extraction, respectively (Table 3). An underestimation of methionine may be due to oxidation during the SDS-PAGE [31] and acid hydrolysis [20]. A high level of glycine was due to contamination from SDS-PAGE. However, these were much lower average % errors compared with the results from other studies (average % error=26.7 [18]) in ABRF trials where proteins are spotted onto PVDF membranes. It is more difficult to analyse electroblotted proteins onto PVDF from SDS-gel due to high contaminations. Fig. 1 is a plot comparing amino-acid compositions of β-lactoglobulin from the two methods, along with the true composition. We conclude that a one-step extraction procedure is adequate to use with PVDF-bound proteins, using Fmoc derivatisation. The TEA mixture (pH 10.0) neutralisation step, necessary for some derivatisation systems, does not improve the quality of the Fmoc-based amino-acid analysis reported here.

3.3. Storage of protein hydrolysates

The stability of protein hydrolysates of β-lactoglobulin (non-PVDF-bound) was tested to determine if hydrolysis could be done significantly prior to analysis. This becomes an important issue when high sample throughput is concerned, as each amino-acid analysis machine has a limited sample capacity per day and hydrolysis is done in batches. Protein hydrolysates were suspended into 250 m*M* borate buffer (pH 8.8) and aliquoted into autosampler vials. Four treatments were examined: storage dry or in solution; 4°C or −20°C. Samples were examined over a period of 21 days. Fig. 2a is the plot of total pmol of amino acids of β-lactoglobulin hydrolysate stored at −20°C in borate buffer or dry, both for up to 14 days. Fig. 2b is the plot of total pmol of amino acids of β-lactoglobulin hydrolysate stored at 4°C or −20°C, both in borate buffer for up to 21 days. Decreasing amounts of total amino acids plateauing to about 50% of immediate assay values were

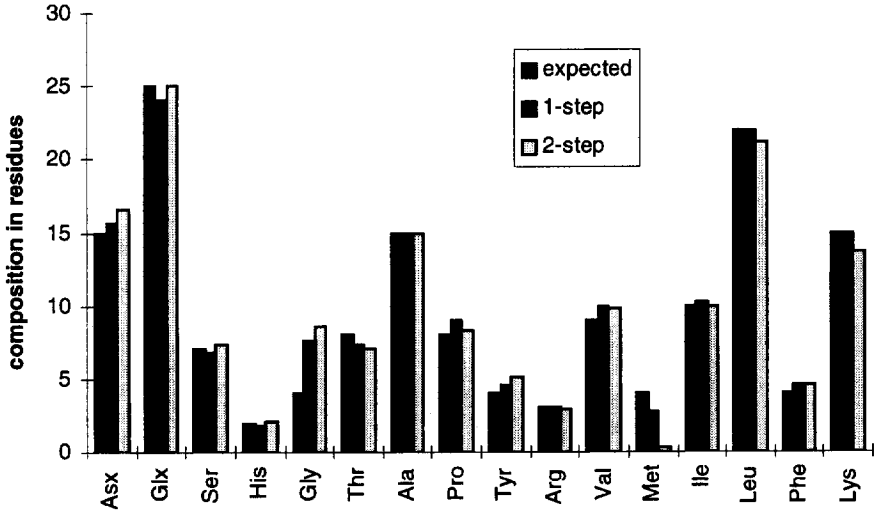

Fig. 1. Amino-acid composition of PVDF-bound β-lactoglobulin after one-step and two-step extraction methods (normalised to alanine).

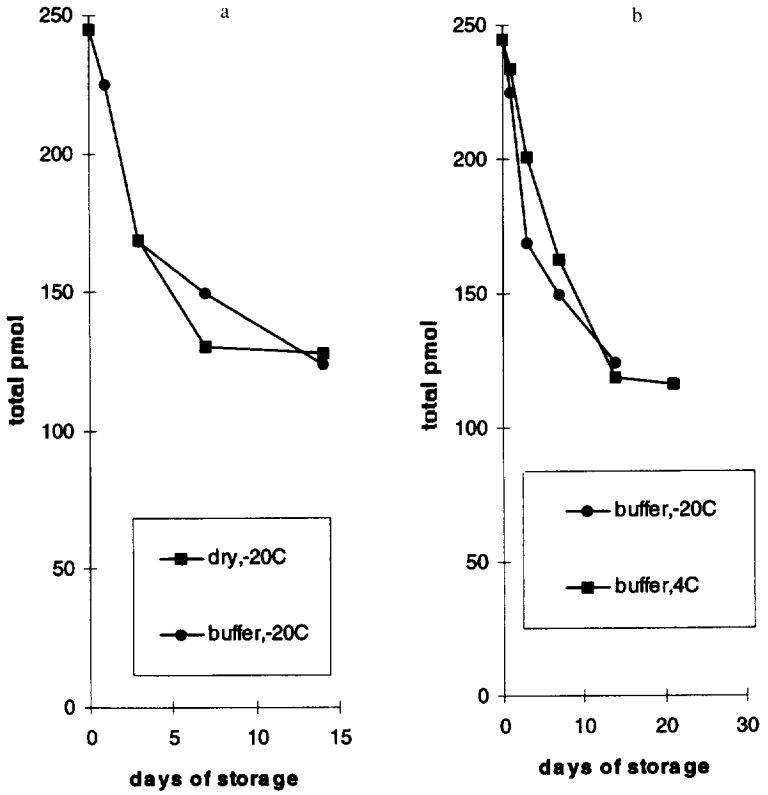

Fig. 2. Total pmol of amino acids from non-PVDF-bound β-lactoglobulin hydrolysate stored (a) at −20°C, in borate buffer, pH 8.8, or dried; (b) in borate buffer, pH 8.8, at 4°C or −20°C.

observed in all cases. However, the amino-acid composition remained similar and close to the expected data throughout the storage period. Fig. 3 shows the amino-acid composition of β-lactoglobulin after different storage times up to 21 days. The mean average % error from day 1 (15.5) to 21 (6.2) was found to be 8.9. Hence, storage does not affect particular amino acids, and the loss of amino acids is independent of temperature between 4 and −20°C. Only lysine showed significant compositional changes during storage, with low recovery immediately after hydrolysis. This improved with time of storage. This experiment has been replicated, and the same results were achieved (data not shown here). We conclude that it is not recommended for samples to be stored for a long period of time if they are to be quantitated. If samples are being analysed qualitatively, they can be stored, although the total yield may be reduced by up to 50%.

3.4. Identification of standard proteins based on their amino-acid composition

The amino-acid compositions of the standard proteins used in the experiments involved with determining the acid concentration for hydrolysis, the extraction method, and protein hydrolysate stability were matched against SWISS-PROT database entries of species of interest used without p*I* and M_r windows. In all 12 cases, proteins were correctly identified. The matching score from rank 1 of each protein is presented in Table 2. As such, it can be concluded that this amino-acid analysis procedure is of high quality, and fulfils the requirements for large-scale proteome studies. A point of particular interest is that we observed that the quality of the amino-acid analysis (based on average % error) was highly correlated to the quality of the score from database searching (Table 3) (correlation coefficient=0.911, $P<0.001$). This demonstrates that accurate amino-acid composition is critical for confident protein identification.

3.5. Applications for identification of proteins from 2-D gel spots

The application of amino-acid analysis for identification of unknown proteins from 2-D electrophoresis was illustrated by Jungblut et al. [10,11] and recently in our laboratory [1,2,8]. Amino-acid composition, M_r and p*I* offer distinctive features of the unknown protein, assisting in finding its homologue in the

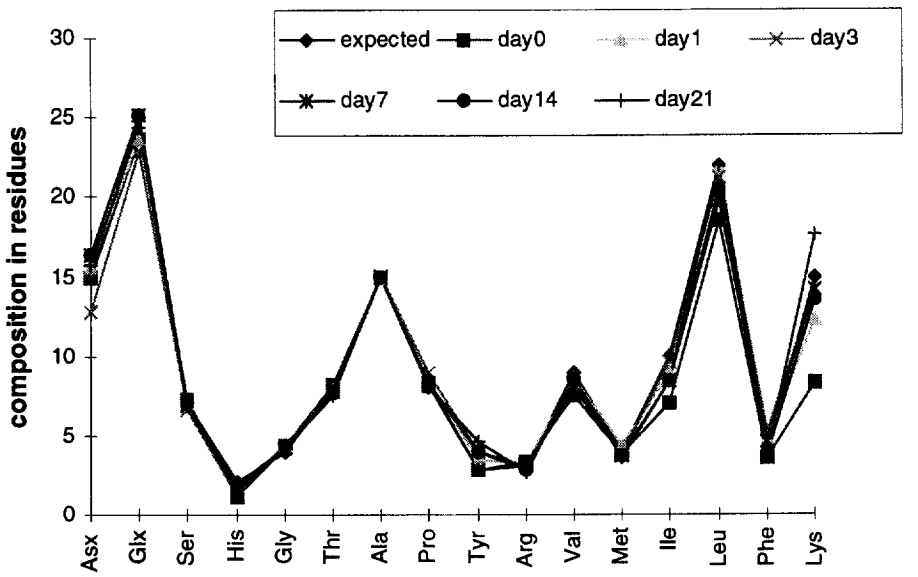

Fig. 3. Amino-acid composition of non-PVDF β-lactoglobulin hydrolysate after different storage times, stored in borate buffer (pH 8.8) at −20°C.

protein databases [1,9]. Demonstrated here is duplicate analysis of three individual protein spots separated by 2-D gel electrophoresis, blotted on PVDF and analysed by amino-acid analysis. These proteins were BSA, human natural killer cell enhancing factor B (NKFB) and human hemoglobin α chain (HBA), as determined by N-terminal sequencing from another 2-D gel. Using experimental amino-acid composition data and pI and M_r (estimated from the 2-D gel), the correct identification was obtained by matching against all the entries of the SWISS-PROT database for the particular species. Rank 1, with the lowest score, gave the correct identification. Table 4 shows the matching data for amino-acid analysis alone from ExPASy [32] World Wide Web server by specified species. The patterns of score and rank give high confidence of correct identification for this single species matching [1]. When the estimated pI and M_r values were included in the search, the degree of confidence increased (Table 5) in all cases but one, the score for the second ranked protein increased significantly. This indicates that the importance of 2-D gel electrophoresis is not only in providing a high-resolution separation of complex protein mixtures, but also in providing estimates of the pI and M_r, which assists in correct and unambiguous identification [1].

4. Conclusion

Amino-acid analysis methods described here provide highly reproducible techniques for identifying proteins separated by 1-D and 2-D gel electrophoresis. Minimum sample handing allows high-sensitivity analysis of highly pure protein spots separated by 2-D gel electrophoresis for the purpose of large-scale proteome studies.

Table 4
Matching data of the closest SWISS-PROT entries (first three ranks) for the specified species based on amino-acid composition alone

Rank	Score	Protein entry	(pI	M_r)	Description
The closest SWISS-PROT entries for the species Human:					
SpotNb hba sample1					
1	5	HBA-Human	8.73	15 126	Hemoglobin α-chain
2	54	HBE-Human	8.69	16 072	Hemoglobin ε-chain
3	62	HBB-Human	6.81	15 867	Hemoglobin β-chain
SpotNb hba sample2					
1	54	HBA-Human	8.73	15 126	Hemoglobin α-chain
2	71	SSR5-Human	9.58	39 202	Somatostatin receptor type 5
3	72	SSR4-Human	9.09	41 894	Somatostatin receptor type 4
The closest SWISS-PROT entries for the species Bovin:					
SpotNb bsa sample1					
1	5	ALBU-Bovin	5.56	66 409	Serum albumin
2	22	CNRC-Bovin	5.54	98 797	Cone CGMP-specific 3′,5′-cyclic
3	27	PIP1-Bovin	5.86	138 714	1-Phosphatidylinositol-4,5-bisphosphate
SpotNb bsa sample2					
1	7	ALBU-Bovin	5.56	66 409	Serum albumin
2	22	CNRC-Bovin	5.54	98 797	cone CGMP-specific 3′,5′-cyclic
3	24	PLC1-Bovin	6.42	23 022	Placental lactogen
The closest SWISS-PROT entries for the species Human:					
SpotNb nkfb sample1					
1	5	NKFB-Human	5.66	21 891	Natural killer cell enhancing factor B
2	11	TSA-Human	6.83	21 857	Thiol-specific antioxidant protein (PRP)
3	22	AATC-Human	6.81	46 135	Aspartate aminotransferase, cytoplasmic
SpotNb nkfb sample2					
1	6	NKFB-Human	5.66	21 891	Natural killer cell enhancing factor B
2	18	TSA-Human	6.83	21 857	Thiol-specific antioxidant protein (PRP)
3	23	LIPP-Human	6.29	49 519	Lipase, pancreatic

Table 5
Matching data of the closest SWISS-PROT entries (first three ranks) having p*I* and M_r values in the specified range

Rank	Score	Protein entry	p*I*	M_r	Description
The closest SWISS-PROT entries for the species Human[a]:					
SpotNb hba sample1					
1	5	HBA-Human	8.73	15 126	Hemoglobin α-chain
2	54	HBE-Human	8.69	16 072	Hemoglobin ε-chain
3	73	HBD-Human	7.97	15 924	Hemoglobin δ-chain
SpotNb hba sample2					
1	54	HBA-Human	8.73	15 126	Hemoglobin α-chain
2	86	HBE-Human	8.69	16 072	Hemoglobin ε-chain
3	86	HBAZ-Human	7.92	15 506	Hemoglobin ζ-chain
The closest SWISS-PROT entries for the species Bovin[b]:					
SpotNb bsa sample1					
1	5	ALBU-Bovin	5.56	66 409	Serum albumin
2	53	CNCN-Bovin	5.28	61 006	63 kDa calcium/calmodulin-dependent 3′5′
3	55	PIP6-Bovin	5.87	78 685	1-Phosphatidylinositol-4,5-bisphosphate
SpotNb bsa sample2					
1	7	ALBU-Bovin	5.56	66 409	Serum albumin
2	44	PIP6-Bovin	5.87	78 685	1-Phosphatidylinositol-4,5-bisphosphate
3	45	CNCN-Bovin	5.28	61 006	63 kDa calcium/calmodulin-dependent 3′5′
The closest SWISS-PROT entries for the species Human[c]:					
SpotNb nkfb sample1					
1	5	NKFB-Human	5.66	21 891	Natural killer cell enhancing factor B
2	49	HA2Z-Human	5.44	25 018	HLA class-II histocompatibility antigen
3	58	IL1X-Human	5.46	17 126	Interleukin-1 receptor antagonistprotein
SpotNb nkfb sample2					
1	6	NKFB-Human	5.66	21 891	Natural killer cell enhancing factor B
2	54	RETB-Human	5.27	21 071	Plasma retinol-binding protein
3	58	GRB2-Human	5.89	25 206	Growth factor receptor-bound protein 2

[a] p*I* (8.88; range 7.88–9.88) and M_r (13 540; range 9478–17 602) were estimated from 2-D gel.
[b] p*I* (5.55; range 5.05–6.05) and M_r (67 000; range 53 600–80 400) were estimated from 2-D gel.
[c] p*I* (5.40; range 4.90–5.90) and M_r (21 200; range 16 960–25 440) were estimated from 2-D gel.

Acknowledgments

We acknowledge support for our research on amino-acid analysis by grants from the Australia Research Council, Macquarie University Research grants and GBC Scientific Equipment (Dandenong, Victoria 3175, Australia). MRW is the recipient of an Australian Postgraduate Research Award. J.X.Y. is a MUCAB Ph.D. Scholar.

References

[1] M.R. Wilkins, C. Pasquali, R.D. Appel, K. Ou, O. Golaz, J.-C. Sanchez, J.X. Yan, A.A. Gooley, G. Hughes, I. Humphery-Smith, K.L. Williams and D.F. Hochstrasser, Bio/Technology, 14 (1996) 61.
[2] V.C. Wasinger, S.J. Cordwell, A. Cerpa-Poljak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W. Duncan, R. Harris, K.L. Williams and I. Humphery-Smith, Electrophoresis, 16 (1995) 1090–1092.
[3] J.E. Celis, H.H. Rasmussen, E. Olsen, P. Madsen, H. Leffers, B. Honore, K. Dejgaard, P. Gromov, H.J. Hoffmann, M. Nielsen, A. Vassilev, O. Vintermyr, J. Hao, A. Celis, B. Basse, J. Lauridsen, G.P. Ratz, A.H. Andersen, E. Walbum, I. Kjaergaard, M. Puype, J. Van Damme and J. Vandekerkhove, Electrophoresis, 14 (1993) 1091.
[4] J.I. Garrels and B.R. Franza, J. Biol. Chem., 264 (1989) 5283.
[5] C. Eckerskorn and F. Lottspeich, Electrophoresis, 11 (1990) 554.
[6] J.I. Garrels, B. Futcher, R. Kobayashi, I. Latter, B. Schwender, T. Volpe, J.R. Warner and C.S. McLaughlin, Electrophoresis, 15 (1994) 1466.
[7] P.F. Lemkin and E.P. Letser, Electrophoresis, 10 (1989) 122.
[8] S. Cordwell, M.R. Wilkins, A. Cerpa-Polik, A.A. Gooley, M.W. Duncan, K.L. Williams and I. Humphery-Smith, Electrophoresis, 16 (1995) 438.

[9] U. Hobohm, T. Houthaeve and C. Sander, Anal. Biochem., 222 (1994) 202.
[10] P. Jungblut, M. Dzionara, J. Klose, and B. Wittmann-Leibold, J. Protein Chem., 11 (1992) 603.
[11] P. Jungblut, A. Otto, E. Zeindl-Eberhart, K.-P. PleiBner, M. Knecht, V. Regitz-Zagrosek, E. Fleck and B. Wittmann-Liebold, Electrophoresis, 15 (1994) 685.
[12] S. Shaw, Proc. Natl. Acad. Sci. USA, 90 (1993) 5138.
[13] P.R. Sibbald, H. Sommerfeldt and P. Argos, Anal. Biochem., 198 (1991) 330.
[14] B. Bjellqvist, J.-C. Sanchez, C. Pasquali, F. Ravier, N. Paquet, S. Frutiger, G.J. Hughes and D.F. Hochstrasser, Electrophoresis, 14 (1993) 1375.
[15] B. Bjellqvist, G.J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J.-C. Sanchez, S. Frutiger and D.F. Hochstrasser, Electrophoresis, 14 (1993) 1023.
[16] K.U. Yuksel, T.T. Anderson, I. Apostol, J.W. Fox, R.J. Paxton and D.J. Strydom, in J.W. Crabb (Editor), Techniques in Protein Chemistry IV, Academic Press, San Diego, CA, 1994, p. 185.
[17] F. Gharahdaghi, D. Atherton, M. DeMott and S.M. Mische, in R.H. Augeletti (Editor), Techniques in Protein Chemistry III, Academic Press, San Diego, CA, 1992, p. 249.
[18] G.E. Tarr, R.J. Paxton, Y.-C.E. Pan, L.H. Ericsson, in R.H. Augeletti (Editor), Techniques in Protein Chemistry II, Academic Press, San Diego, CA, 1991, p. 261.
[19] M. Ploug, A.L. Jensen and V. Barkholt, Anal. Biochem., 181 (1989) 33.
[20] G.I. Tous, J.L. Fausnaugh, O. Akinyosoye, H. Lackland, P. Winter-Cash, F.J. Vitorica and S. Stein, Anal. Biochem., 179 (1989) 50.
[21] S. Nakagawa and T. Fukuda, Anal. Biochem., 181 (1989) 75.
[22] U.K. Laemmli, Nature, 270 (1970) 680.
[23] P. Matsudaira, J. Biol. Chem., 21 (1987) 10035.
[24] J.-C. Sanchez, F. Ravier, C. Pasquali, S. Frutiger, N. Paquet, B. Bjellqvist, D.F. Hochstrasser and G.J. Hughes, Electrophoresis, 13 (1992) 715.
[25] H.E. Meyer, E. Hoffmann-Posorske and L.M.G. Heilmeyer, Jr., Methods Enzymol., 201 (1991) 169.
[26] K. Ou, M.R. Wilkins, J.X. Yan, A.A. Gooley, Y. Fung, D. Sheumack and K.L. Williams, J. Chromatogr. A, 723 (1996) 219.
[27] P.A. Haynes, D. Sheumack, J. Kibby and J.W. Redmond, J. Chromatogr., 540 (1991) 177.
[28] P.A. Haynes, D. Sheumack, L.G. Greig, J. Kibby and J.W. Redmond, J. Chromatogr., 588 (1991) 107.
[29] L.B. Smillie and M. Nattriss, in C.T. Mant and R.S. Hodges (Editors), High-performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation, CRS Press, 1991, p. 847.
[30] A. Tsugita and J.J. Scheffler, Eur. J. Biochem., 124 (1982) 585.
[31] K.H. Fantes and I.G.S. Furminger, Nature, 215 (1967) 750.
[32] R.D. Appel, A. Bairoch and D.F. Hochstrasser, Trends Biochem. Sci., 19 (1994) 258.