ELSEVIER

Journal of Chromatography B, 771 (2002) 197–209

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Review

Proteomic map and database of lymphoblastoid proteins

Michel Caron[a,b,*], Naima Imam-Sghiouar[a], Florence Poirier[a], Jean-Pierre Le Caër[c], Valerie Labas[c], Raymonde Joubert-Caron[a]

[a]*Université Paris 13, UFR SMBHI Leonard de Vinci, Bobigny, France*
[b]*Laboratoire Central de Biochimie, Hôpital Avicenne, Bobigny, France*
[c]*Laboratoire de Neurobiologie, Ecole Supérieure de Physique Chimie Industrielle de la Ville de Paris (ESPCJ), Paris, France*

Abstract

Advances in genomics have led to the accumulation of an unprecedented amount of data, giving rise to a new field in biochemistry, proteomics. We used a combination of two dimensional gel electrophoresis, analysis and annotation using third-generation software, and mass spectrometry to establish the proteome maps of lymphoblastoid B-cells, a prerequisite for analysis of drug effects and lymphocyte cell diseases. About 1200 protein spots were detected and characterised in terms of their isoelectric point, molecular mass and expression. The present status of proteomic technologies, as well as a description of the usefulness of human hematopoietic cells proteomic database are discussed. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Reviews; Proteomics; Lymphoblastoid proteins

Contents

1. Introduction 198
2. Background: survey of applicability of proteomics on lymphoid cells 198
3. The basic steps of lymphoblastoid proteome analysis 199
   3.1. Two-DE of soluble lymphoblastoid cell proteins 199
   3.2. Computer-assisted two-dimensional gel electrophoresis approach for studying variations in protein expression (differential proteomics) 202
   3.3. Lymphoblastoid protein identification 202
   3.4. The BPP human hematopoietic cell lines proteomic database 203
4. Analysis of cell signalling by combined precipitation and electrophoresis techniques 203
5. Phospho-proteomics in lymphoid cells 206
6. Perspectives of proteomics for biomedical applications in immunology and haematology 207
Acknowledgements 208
References 208

*Corresponding author. Biochimie des Protéines et Protéomique, UER SMBH-Léonard de Vinci, Université Paris 13, 74, rue Marcel Cachin, 93017 Bobigny, France. Tel.: +33-1-4838-7754; fax: +33-1-4838-7777.
E-mail address: caron@smbh.univ-paris13.fr (M. Caron).

1570-0232/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1570-0232(02)00040-5

1. Introduction

The availability of complete genome sequences necessitates the development of functional assays to analyse the hundreds of thousands of predicted gene products. Such approaches are collectively referred to as 'functional genomics'. Proteins, not genes, however sustain function. Furthermore, the amount of mRNA in a cell often shows no correlation with the amount of protein that gets produced by the cell [1]. Proteins can undergo innumerable slight chemical changes that can profoundly alter their activity. Molecular biology techniques have dominated the functional analysis of gene products for many years but protein analysis is coming back into focus because it is now widely realised that if we want to know what is happening to a cell's proteins, we have to study the proteins themselves. Moreover, the release of a complete draft of the human genome [2] allows, in principle, identification of all human gene products via proteomics methods. Now, we face the challenge of making use of this information for improving healthcare and discovering markers for diseases or new drug targets. In this context, proteome analysis technology is used extensively in conducting discovery research of biology and has become one of the most essential technologies in functional genomics.

Electrophoresis methods represent an essential tool for the separation of proteins at the laboratory scale. Among them, high-resolution two-dimensional electrophoresis (2-DE) is the prevailing technique for analysis of the whole protein expression profile of a given cell type, under a set of specific conditions [3,4]. Image analysis software complete the process by allowing spot quantification and comparison [5,6]. Mass spectrometry methods represent the methods of choice for the identification of proteins in proteomics projects with a good throughput. Typically, spots on the 2-DE gel are excised for identification using mass spectrometry (e.g. MALDI-TOF) coupled to protein sequence database searching. In this context, proteome expression profiles of well-documented human cell lines offer insight into the functional role of specific proteins [7]. Different expression profiles can be compared by subtractive analysis, and facilitate investigation on differentiation or programmed cell death specific proteins, on markers in diagnosis, and on drug target in disease treatments (differential proteomics). The obtained information should finally be stored in 2-DE databases enhancing the efficiency of proteome research [8].

2. Background: survey of applicability of proteomics on lymphoid cells

The term 'proteome' refers to the total PROTeins expressed by a genOME and is the systematic analysis of protein profile of a given cell or tissue type [9]. Currently, analytical efforts in proteomics are focusing on the identification of proteins that play a role in a particular biological scenario. Proteomics is based on a variety of methods allowing one to isolate, detect, identify, and to analyse quantitatively thousands of proteins comprising the cellular proteome. Proteome research basically involves the following experimental steps: (i) separation of proteins in the proteome; (ii) identification/characterisation of individual proteins; and (iii) storage of the information in databases.

To date, 2-DE is the key proteomic technology. Due to its unique resolution and ability to separate thousands of proteins from one single sample, it is currently the method of choice to analyse simultaneously the protein composition of cells and to reveal quantitative and qualitative changes in proteins modulating multiples cell functions. This technique combines isoelectric focusing (TEF) in the first dimension with sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS–PAGE) in the second dimension [10], and is capable of separating several thousands of proteins on a single 2-DE gel.

Mass spectrometry, in combination with advanced bioinformatics, has now made it possible to identify proteins present in small amounts on a 2-DE. Therefore, identification of 2-DE resolved proteins in now primarily done using mass spectrometric techniques, such as matrix-assisted laser desorption ionisation (MALDI) or electrospray ionisation (ESI) mass spectrometry. The combination of 2-DE and MALDI or ESI spectrometry has the potential to resolve a complex proteome and characterise its individual proteins with respect to their identities, quantities, and post-translational modifications [11,12].

A human lymphoid cell extract must be regarded

as a raw biological material of extreme complexity. It consists of numerous proteins which are present at very different concentrations. To our knowledge, S.M. Hanash and colleagues have undertaken the first effort aimed at developing a database of lymphoid proteins about 8 years ago, at the University of Michigan School of Medicine [13]. This database contained 2-DE patterns and derived information pertaining to polypeptides of unstimulated and stimulated mature T cells and immature thymocytes, cultured T cell lines, and a variety of other relevant cell populations. Several markers have been studied in detail, including stathmin, a cytosolic phosphoprotein implicated in signal transduction [14,15]. The database was originally intended as a resource for biologists. However, the operation of the database was sufficiently complex to inhibit its use by biologists with limited familiarity with computer analysis. More recently, additional resources have been put into place to mine more effectively data pertaining to protein expression in lymphoid cells, and to move existing data to a user-friendly interface by using Web Browsers [16].

More specifically, B. Wittmann-Liebold, K. Bommert, and colleagues identified apoptosis-associated proteins in the human Burkitt lymphoma cell line BL60 [17]. After anti-IgM treatment and separation of apoptosis-committed cells, comparison of the 2-DE protein patterns from apoptotic and non-apoptotic cells showed differences in approximately 80 spots. Subsequent analysis by MALDI-TOF identified D4-GDI as a protease substrate during anti-IgM-mediated apoptosis [18]. Recently, a 2-DE database of the human leukaemia Jurkat T-cell line was published by T. Rudel and colleagues [19], available at http://www.mpiib-berlin.mpg.de/2D-PAGE/. In 2000, peptide mass fingerprinting has identified 67 protein spots of this 2-DE map.

In recent years we have undertaken an effort to identify and develop a database of lymphoid proteins detectable by 2-DE, and more particularly of soluble proteins in human lymphoid B-cell lines, that constitute widely used models for investigations on B cells. In addition, the database contains 2-DE maps of leukaemia (TF1, K562, KG1a) and promyelocytic (HL60) cell models [20]. The identification of proteins on 2-DE maps was obtained by various techniques including gel comparison (matching), peptide mass fingerprinting using MALDI-TOF, and identification by tandem mass spectrometry using LC–MS–MS. We used proteomics to separate complex protein mixtures of hematopoietic cells into their individual components, to compare expression profiles of sample pairs, and to determine the global protein response of a cell to a given set of conditions.

By way of introduction, the main components of the proteomic technology as our laboratories carry it out will be described in outline (Fig. 1). Briefly, the proteins extracted from cultured human cell lines are separated by 2-DE with immobilised pH gradient isoelectric focusing (IPG-IEF) in the first dimension and SDS–PAGE in the second dimension (IPG-DALT) [21,22]. The proteins are detected using silver staining, and the resulting images analyzed using 2-DE analysis software [7,22]. After statistical analysis, features of interest are chosen for cutting and submitted to mass spectrometric analysis. This allows the identification (annotation) of the protein features, and allows constituting reference maps to evaluate the significance of protein changes observed in the context of the experiment or disease [23,24]. This strategy will be further described in the context of lymphoblastoid proteins analysis. Lymphoblastoid cell lines are a suitable model for investigations on B cells, and for biomedical applications in immunology and haematology.

3. The basic steps of lymphoblastoid proteome analysis

3.1. Two-DE of soluble lymphoblastoid cell proteins

Proteins were extracted from lymphoblastoid B cells as previously described in a buffer containing 50 m*M* Tris pH 7.4, 10 m*M* EDTA, 65 m*M* DTT, 1.5 m*M* PMSF, and anti-proteases (3×10^8 cells/ml) [7]. After centrifugation, the clear supernatant was supplemented with 7 *M* urea, 2 *M* thiourea, and 4% CHAPS. The proteins were then separated by 2-DE as described [7,22,23]. Briefly, samples were applied overnight to 17 or 18 cm long Immobiline Dry Strips, using IPG gradients: 4–7 (Amersham Phar-

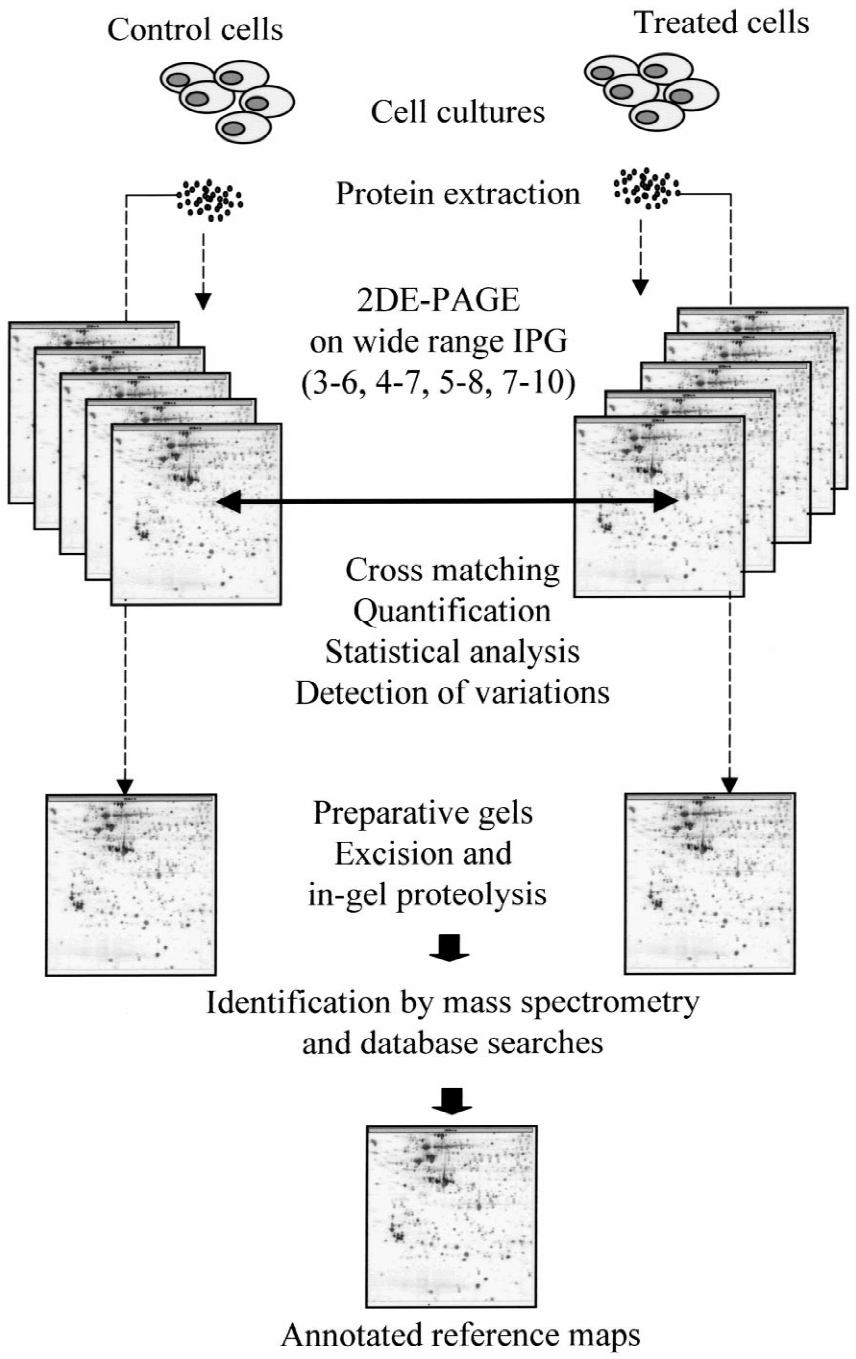

Fig. 1. Steps involved in analysing lymphoblastoid cell cultures by proteomics.

macia Biotech, Uppsala, Sweden), 3–6, 5–8, or 7–10 (Bio-Rad, Richmond, CA, USA). A total of 100 or 300 μg of protein was applied for analytical and preparative 2-DE gels, respectively. IPGphor (Amersham Pharmacia Biotech, Uppsala, Sweden) or Protean JEF Cell (Bio-Rad, Richmond, CA, USA) were used for the first dimension. Focusing was carried out for a total of around 45 000 Vh for analytical samples and 90 000 Vh for preparative samples, depending on the IPG used. After focusing, the gels were equilibrated for 20 mm in a buffer containing 50 m*M* Tris–HCl, pH 8.6, 6 *M* urea, 30% glycerol, 65 m*M* DTT, 1% SDS, and a trace of bromophenol blue. A second equilibration of 30 min was carried out with 53 m*M* iodoacetamide instead of DTT. Protean II xi Cell 10 (Bio-Rad, Richmond, CA, USA) was used for the second dimension (gel size 18 cm×20 cm). Fig. 2 shows the analytical silver-stained 2-DE reference map of the lymphoblastoid PRI cells (IPG gradient: 4–7). Expression patterns or more than 1200 protein spots were detected and observed through the entire series of gels. Protein

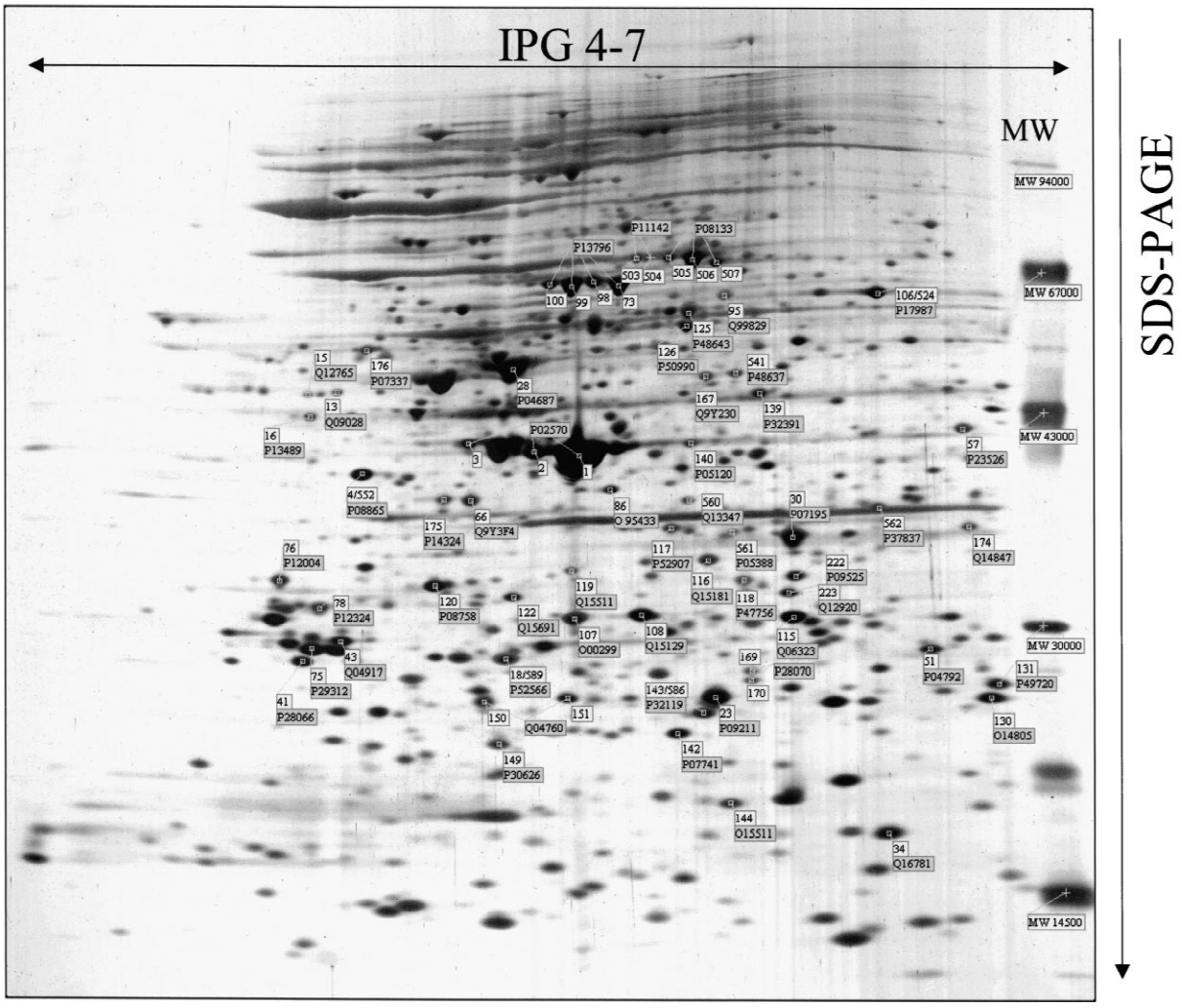

Fig. 2. 2-DE map of protein extracted form PRI cell line. Proteins were separated on a linear pH 4–7 IPG strip, followed by an 8–18% SDS–PAGE. The gel was stained with silver nitrate. The proteins identified are designated with their landmarks (spot numbers) and their accession numbers in SWISS-PROT. The protein names are listed in Table 2.

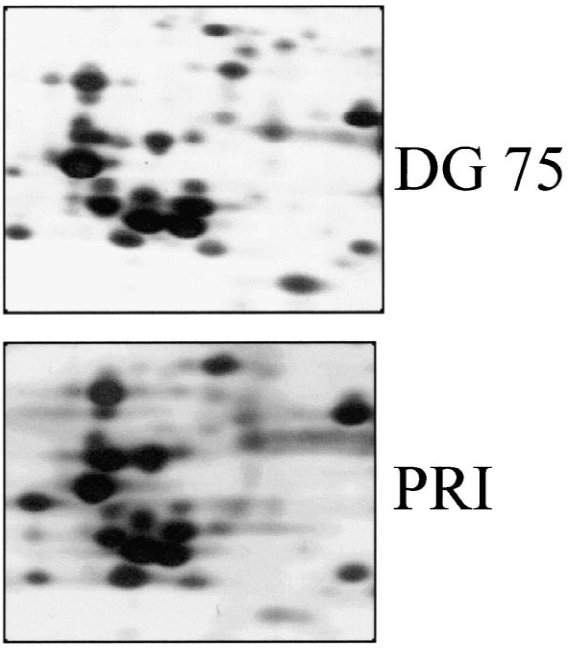

Fig. 3. Enlargements of a portion of the 2-DE of protein extracts from two lymphoid B cell lines: DG75, a Burkitt lymphoma cell line, and PRI, a lymphoblastoid cell line.

spots could easily be cross-matched to each other, even between different lymphoid B cell lines, indicating the reproducibility of the method. This is illustrated in Fig. 3, that compares the same window of 2-DE obtained on the Burkitt lymphoma line DG75 and on the lymphoblastoid line PRI.

3.2. Computer-assisted two-dimensional gel electrophoresis approach for studying variations in protein expression (differential proteomics)

In differential proteomics, two (or more) proteomes are compared by a subtractive analysis [25] in which differences due to drug treatment, culture conditions, or diseases can be observed. The result is a number of protein spots, which are up- or down-regulated in the proteome. For the proteomes of lymphoid cells, silver-stained gels were scanned using a GS-700 densitometer 10 (Bio-Rad, Richmond, CA, USA). The computer image analysis was carried out using MELANIE 2.3 and MELANIE 3.5 software (GeneBio, Geneva, Switzerland), allowing automatic detection and quantification of protein spots, as well as the re-sizing, alignment and matching between our different 2-DE [26]. A calibration in p*I* and M_r was done for each gel. Spot volume of each protein was computed by considering spot area and spot optical density. The quantification of each spot of interest was expressed as percent volume (%*V*), where: %*V*, spot volume/volumes of all spots of interest resolved in the gel.

3.3. Lymphoblastoid protein identification

For protein identification by mass spectrometry, colloidal Coomassie Brilliant Blue- or silver-stained preparative 2-DE gels were matched against analytical gels using MELANIE software. Protein spots with a wide range of molecular masses (M_r=15–150), isoelectric points (4.5–7.5), and abundances were subjected to MS protein identification. Excised stained protein spots were digested with bovine trypsin as described [27] with minor modifications. Briefly, gel pieces derived from 2-DE are washed several times with 50% acetonitrile. A final washing step with acetonitrile leads to a shrinking of the gel piece, which is swollen in 10 m*M* DTT, 100 m*M* NH_4HCO_3 for 45 min, followed by 55 m*M* iodoacetamide, 100 m*M* NH_4HCO_3 for 45 min. Gel pieces are then dried another time and reswollen with 30 μl protease solution (trypsin at 0.1 mg/ml in 50 m*M* NH_4HCO_3). Digestion is performed overnight at 37 °C. After centrifugation, the supernatant is collected the resulting peptide mixture is concentrated and desalted using ZipTip pipette tips (Millipore, Bedford, CA, USA).

A peptide mass profile was obtained by using an Applied Biosystems MALDI-TOF Voyager-DE Mass Spectrometer. The peptide masses obtained were used for database searches for protein identification. A protein can be identified by peptide mass fingerprinting if its mass fingerprint fits one (and ideally only one) of the fingerprints computed from a known sequence. The proteins were identified based on their peptide mass map using the database search programs PeptIdent (developed by the Swiss Institute of Bioinformatics) [28] and ProFound (developed by the Rockefeller University) [29]. The search parameters are shown in Table 1. Our success rate of identification using the mass fingerprint spectra varies between 60 and 70%. In most cases the

Table 1
The search parameters in PeptIdent and ProFound

Parameter	Parameter setting
Species searched	Homo sapiens
Digest used	Trypsin
Peptide mass accuracy	0.1 Da
Partial modification	Methionines oxidized
Cysteine modification	Carbamidomethylation
Peptide masses are	Monoisotopic
Charge state	MIH+
Number of uncleavaged sites	1

Database searched: PeptIdent: SWISS-PROT+TrEMBL; ProFound: NCBInr.

peptides identified for a given protein were derived from regions along the entire length of the protein, indicating the observed polypeptides were not the result of proteolytic degradations. When there was no unambiguous fit (e.g. multiple candidate proteins), a combination of fingerprint information and of sequence information from additional tandem mass spectrometry experiments (LC–MS–MS) was used. LC–MS–MS experiments (Micromass' Q-Tof) were used to identify the protein in the protein sequence database of NCBI using the Mascot program [30] for processing of protein sequences. For the lymphoblastoid B cell line PRI, comparing the theoretical value of the molecular mass and p*I* of each protein to that of the observed value, most of the proteins (60% of the analyzed spots) were identified in the expected position on the gel ($\Delta pI<0.2$, $\Delta M_r<20\%$). Some of the other proteins with strong matches migrated to a somewhat unexpected p*I* (Elongation factor 2 with three experimental p*I* of 6.97, 7.10, and 7.23, for a theoretical value of 6.41), due probably to the addition of phosphoryl groups on the polypeptide chain. Twenty-five percent of the spots gave clear peptide peaks on mass spectrometry but did not match any known polypeptide. These results are illustrated by a list compiling proteins identified in PRJ cells in the range of p*I* 4–7, and are reproduced in the Table .

3.4. The BPP human hematopoietic cell lines proteomic database

This study allowed the construction of new annotated reference maps included in the BPP 2-DE database. This database has been described in details previously [20]. Reference maps for the following cellular types are available: leukaemia cells, promyelocytic cells, lymphoma cells, and lymphoblastoid cells. Spots flagged with a label correspond to identified polypeptides. A view of any map with labels of identified proteins is obtained by clicking on the gel icon, while a view of the gel without labels is obtained by clicking on 'without index' inside the window (Fig. 4).

4. Analysis of cell signalling by combined precipitation and electrophoresis techniques

Even if thousands of proteins are separated and may be visualised on a 2-DE gel, these proteins represent only a part of the proteome [31]. In fact, it is generally considered that the greater part of the protein mass found within leaving cells represents about 10% of the gene products. Therefore, only a fraction of the cellular proteins are present at levels sufficient for analysis. The current estimation of the human genome being of about thirty thousand genes, taking into account the alternative splicing one may estimate a transcriptome of about 60 000 mRNA, and a proteome of at least 300 000 polypeptides. If we suppose that 1000–2000 spots are separated on a 2-DE corresponding to the same number of polypeptides, and even if we consider that only a portion of the genome is translated in each cell type, the result is far from the theoretical definition of the proteome. An approach to solve this problem is the fractionation of the sample in order, for instance, to concentrate the proteins of interest prior to the first dimension of the gel separation [32,33]. Several approaches have been proposed for the prefractionation of the proteins, that can be separated in:

- protein-by-protein approaches, for elucidating the role of a peculiar protein in complex biological regulation processes [34,35]; and
- analysis of the proteome of a family of proteins, for instance the phospho-proteome, the glycoproteome [36–38].

We have used these two types of approaches, to study modifications of phosphorylation during the

Table 2
Summary of human lymphoblastoid cell proteins identified on the 2-DE gel

Landmark (spot No.)	SWISS-PROT Acc. No.	Protein identified as	p*I* gel, M_r		Δp*I*	ΔM_r (%)	No. of matching peptides	Sequence coverage (%)	Identification
43	Q04917	14-3-3 protein ETA	4.92	28 500	0.16	1.8	5	17.1	MS
75	P29312	14-3-3 protein ZETA/DELTA	4.83	27 900	0.27	11	16	52.7	MS
561	P05388	60 S PROT PO ribosomal	5.59	37 286	nd	nd	nd	nd	Match
4	P08865	67 kDa laminin receptor	4.97	40 536	0.18	18.9	9	29.8	MS
1	P02570	Actin, BETA	5.29	41 899	0.00	0.5	16	46.3	MS
2	P02570	Actin, BETA	5.14	42 072	0.16	0.9	14	38.8	MS
3	P02570	Actin, BETA	5.09	42 196	0.21	1.4	10	24.6	MS
139	P32391	Actin-2	5.68	45 679	0.09	3.9	16	36.5	MS
57	P23526	Adenosyl homocysteinase	6.15	42 596	0.12	12.0	13	21.8	MS
222	P09525	Annexin IV	5.72	33 059	0.13	7.7	12	28.9	MS
120	P08758	Annexin V	5.05	32 300	0.11	10.9	22	62.4	MS
507	P08133	Annexin VI	5.54	78 186	nd	nd	nd	nd	Match
505	P08133	Annexin VI	5.45	80 364	nd	nd	nd	nd	Match
506	P08133	Annexin VI	5.49	76 814	nd	nd	nd	nd	Match
142	P07741	APRT	5.56	22 170	0.23	11.9	6	40.2	MS
144	O15511	AR 16	5.60	18 200	0.13	10.3	4	33.1	MS
118	P47756	CAPB	5.65	32 800	0.04	6.6	10	36.4	MS
117	P52907	CAZI	5.55	36 500	0.10	9.8	14	53.1	MS
107	O00299	Chloride intra channel PROT 1	5.26	30 359	0.24	11.1	7	36.5	MS
95	Q99829	COPINE I	5.62	61 305	0.11	3.7	12	18.8	MS
122	Q15691	EB1	5.06	32 400	0.13	5.2	13	42.2	MS
167	Q9Y230	ECP 51	5.61	50 000	0.12	12.3	15	40.2	MS, MS–MS
8	P56537	EIF-6	4.60	27 000	0.04	40.9	7	44.9	MS
175	P14324	FPP syntethase	5.07	38 600	0.0	5.0	11	28.3	MS
18	P52566	GDIS-HUMAN	5.14	22 590	0.04	15.5	11	59.2	MS
25	P41250	GLYCL-TRNA-SYNTH	6.07	77 774	0.19	0	11	14.8	MS
541	P48637	GSHB	5.60	52 854	nd	nd	nd	nd	Match
23	P09211	GTP-HUMAN	5.61	24 522	0.16	5.2	7	60.8	MS
504	P11142	Heat shock cognate 71 kD	5.62	74 900	nd	nd	nd	nd	Match
503	P11142	Heat shock cognate 71 kD	5.49	73 800	nd	nd	nd	nd	Match
51	P04792	HSP 27 kD	6.06	28 000	1.77	20.3	12	46.2	MS
86	O95433	Hypothetical 38.3 kD PROT	5.53	41 900	0.12	8.7	13	37.0	MS
15	Q12765	Hypothetical protein KIAA0193	4.82	45 274	0.21	13.6	8	24.3	MS
560	Q13347	IF32	5.51	38 381	nd	nd	nd	nd	Match
113	Q06323	IGUP I-5111	5.61	30 400	0.14	14.6	6	24.4	MS
115	Q06323	IGUP I-5111	5.72	30 479	0.03	15.4	17	70.7	MS
116	Q15181	Inorganic pyrophosphatase	5.60	34 200	0.06	4.5	9	33.9	MS, MS–MS
223	Q12920	KI nuclear autoantigene	5.71	32 000	0.42	3.5	7	24.7	MS
99	P13796	L-plastin	5.26	64 000	0.06	11.6	18	29	MS
100	P13796	L-plastin	5.22	64 000	0.02	10.7	13	21.2	MS
98	P13796	L-plastin	5.35	64 000	0.15	9.8	16	31.4	MS
73	P13796	L-plastin	5.48	63 000	0.28	12.13	13	21.5	MS
150	Q04760	Lactoylglutathione lyase	5.10	24 181	0.15	18.3	9	36.1	MS
151	Q04760	Lactoylglutathione lyase	5.25	24 453	0.0	16.0	9	36.1	MS
174	Q14847	LASP-1	6.17	36 815	0.45	18.6	6	29.9	MS
30	P07195	LDH-H chain	5.72	36 000	0.0	2.0	14	37.2	MS
76	P12004	PCNA	4.73	32 800	0.17	12	18	64	MS
39	P35080	Profilin 2	5.93	15 131	0.09	0.3	2	11.5	MS
108	Q15129	Proteasome activator subunit B	5.60	30 600	0.06	10.3	19	66.1	MS
170	P28070	Proteasome BETA chain	5.70	25 790	0.19	5.1	8	45.7	MS

Table 2. Continued

Landmark (spot No.)	SWISS-PROT Acc. No.	Protein identified as	p*I* gel, M_r	Δp*I*	ΔM_r (%)	No. of matching peptides	Sequence coverage (%)	Identification
169	P28070	Proteasome BETA chain	5.66 26 448	0.23	8	5	27.9	MS
131	P49720	Proteasome THETA chain	6.24 25 645	0.02	12	6	28.3	MS
41	P28066	Proteasome ZETA chain	4.80 27 000	0.11	11.8	7	33.5	MS
176	P07337	Protein disulfide isomerase	4.99 52 210	0.31	6.5	15	38.3	MS
13	Q09028	Retinoblastoma binding PROT P48	4.90 45 549	0.16	4.7	9	22.1	MS
16	P13489	Ribonuclease inhibitor	4.82 42 848	0.11	0.3	16	40.4	MS
130	O14805	RNA-binding protein REGUL	6.22 24 522	0.11	18.8	12	55.0	MS
149	P30626	Sorcin proteine 22 kD	5.08 21 370	0.19	0.8	10	56.5	MS
119	Q15511	Spermidine synthase	5.26 33 389	0.04	1.3	8	25.5	MS
106	P17987	T-complex protein 1 ALPHA	5.94 61 860	0.14	2.5	9	21.3	MS, Match
126	P50990	T-complex protein 1 THETA	5.58 58 049	0.09	6.9	20	40.5	MS
125	P48643	TCP-1-EPSILON	5.15 31 712	0.13	2.9	21	46.1	MS
143	P32119	Thioredoxin peroxidase 1	5.59 23 513	0.06	12.4	11	41.9	MS
562	P37837	TRANS aldolase	5.92 37 629	nd	nd	nd	nd	Match
7	P12324	Tropomyosin cytoskeletal type	4.80 30 000	0.05	3.2	8	22.4	MS
78	P12324	Tropomyosin cytoskeletal type	4.85 30 900	0.1	6.3	26	74.6	MS
28	P04687	Tubulin ALPHA-1	5.15 48 838	0.14	3.2	13	41.2	MS
26	P05217	Tubulin BETA-2	5.06 46 809	0.21	6.5	7	12.6	MS
34	Q16781	Ubiquitin–protein ligase	5.96 16 700	0.13	0.8	8	40.8	MS
66	Q9Y3F4	UNR-interacting protein	5.09 38 500	0.11	0.1	13	46.3	MS
140	P05120	Urokinase inhibitor	5.59 42 807	0.14	3.5	14	27.2	MS

Proteins extracted from PRI cell line were separated on a linear pH 4–7 IPG strip, followed by a 8–18% SDS–PAGE. Proteins were identified by peptide mass mapping and protein database searching using PepIdent and Profound software. Some identifications were obtained by matching or by MS–MS using Mascot software.

transduction of a signal in lymphoid cells. Phosphorylation is an important element for regulating the activity of proteins. Most of the known cellular processes involve modulation of cellular signalling pathways via changes in protein phosphorylation. Reversible protein phosphorylation regulates fundamental cellular events, including cell growth, division, differentiation apoptosis, signal transduction [39,40]. Unravelling the complex regulatory networks and signal transduction systems in lymphoid cells is, therefore, a major focus of current cell biochemistry and biomedicine. Consequently, detailed characterisation of phosphoproteins is required to fully understand protein function and regulatory events.

An example that illustrates the protein-by-protein approach is the elucidation of the first step of the cellular signalisation induced by a dimeric ligand on a tyrosine–phosphatase receptor, CD45 [34,41]. CD45 is a highly glycosylated membrane protein, with both extracellular and intracellular domains. The binding of the CD45 ligand, galectin 1, induces its ligation. This ligation blocks the intracellular catalytic site in CD45. In B cells CD45 activity is mediated, at least in part, via control of the activity of a tyrosine kinase, Lyn [42]. CD45 modulates its function by dephosphorylating a phosphotyrosyl residue found at its C-terminus. Our hypothesis was that the binding of galectin 1 to CD45 results in a blockage of its phosphatase activity, resulting in an hyperphosphorylation of Lyn and therefore an inhibition of its kinase activity. To test this hypothesis, Lyn was immunopurified from B lymphoid cells before treatment of the cells by galectin 1 and at different times thereafter. The elution of Lyn from the immunoadsorbant was obtained by incubation in IPG buffer, that allowed submitting directly the eluate to IEF. Then, the isoforms of Lyn were resolved by 2-DE. In proteomic analysis, phosphorylated proteins are generally noticed to shift positions which can be resolved by 2-DE. Such shifts are due to changes in the isoelectric point with little

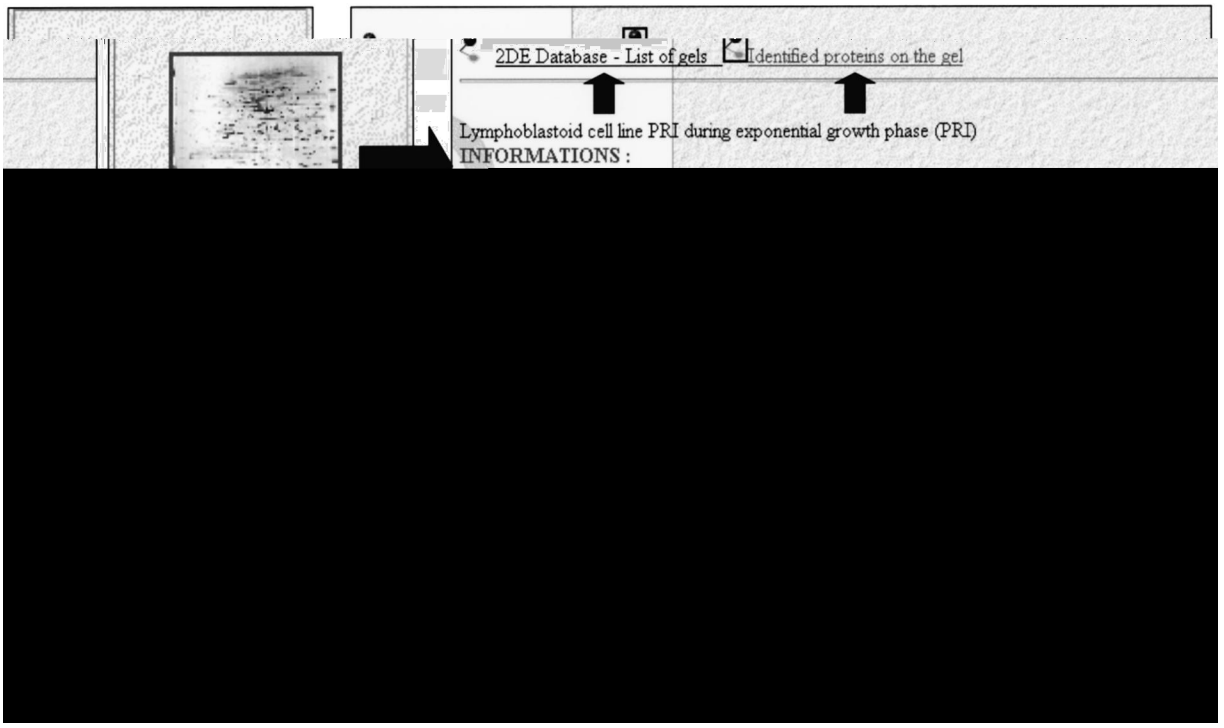

Fig. 4. A partial view of a page of the BPP 2-DE database (http://www-smbh.univparis13.fr/lbtp/Biochemistrv/Biochimie/bciue.htm). Clicking on the gel icon opens a view of the gel (http://www-smbh.univ-paris13.fr/lbtp/Biochemistrv/Biochimie/gelpri4-7.htm).

or no change in apparent molecular mass. After the binding of galectin 1 to CD45, only two spots were modified, these two spots showing a same molecular mass of about 58 kDa and a difference in isoelectric point of 0.3 units corresponding to the addition of a phosphoryl group on the polypeptide. This proteomic study demonstrated that in lymphoblastoid cells the galectin 1-induced ligation of CD45 results in the phosphorylation of one tyrosine in the negative regulatory site of the 58 kDa form of the tyrosine kinase Lyn.

5. Phospho-proteomics in lymphoid cells

Post-translational modifications of proteins are a common means to target proteins, regulate their activities and to mediate communication between proteins and cells. Many times, altered protein profiles of tissues or cells are the result of altered protein modification rather than altered gene expression. Two-DE in combination with MALDI and electrospray MS have been established as powerful tools for investigation, visualisation and identification of phosphoproteins that are present in the cell at a given time [43–45]. In order to study the phosphoproteome of lymphoid cells, we use two different strategies. On one hand, for the affinity separation of the proteins phosphorylated on tyrosine, and of complexes containing these proteins, immunoprecipitations are performed on immobilised anti-phosphotyrosine. The elution is obtained in the buffer used for the first dimension of the 2-DE. On the other hand, global enrichments in phosphorylated proteins are obtained by metal chelate affinity chromatography on ferric ions. To illustrate the results that are obtained, Fig. 5 shows an enlargement area of 2-DE gels obtained without or with immunoprecipitation of tyrosine phosphorylated proteins in lymphoblastoid cells. In the same conditions of separation, about 1200 protein spots were detected on a whole extract of soluble proteins, and on about

insights into important cellular processes and the pathology of diseases. Moreover, proteomics is not just the systematic separation, cataloguing, and study of all the proteins produced in a cell type or an organism. It is also the study of how proteins change structure, interact with other proteins, and ultimately give rise to disease or healthy, even if function has long been treated as a non-contextual attribute of a given protein [8,48]. A thorough understanding of cell regulation and dynamics necessitates not only identification of individual gene products but also investigation of spatial protein distribution, composition of protein complexes and post-translational modifications of proteins. Beyond the identification of proteins involved in a particular physiological situation, many aspects of proteomics require more detailed characterisation of the proteins involved, and the interaction of proteins in a complex network adds up to how an organism functions. Until now, and for practical reasons, proteomics databases represent function by a set of key words that does not translate the fact that the same protein can participate in several distinct cellular processes depending on the global pattern of its molecular interactions. Therefore, functional proteomics will probably play an increasing role in biomedicine in the future.

The key to understanding health and disease within a cell or an organism is, therefore, to understand how its proteins function. This should be largely dependent upon the answers to a number of issues, which includes both technical and medical problems. For such answers, especially regarding the technical problems, the future will depend whether 2-DE will continue to be regarded as the method of reference to perform highly resolute protein separations, or, by contrast, will be replaced at least in part by processes using chromatography. Recently, an approach using affinity purification through an isotope-coded affinity tag (ICAT) was developed for the comparative quantitation of differential protein expression [49]. One may predict that there will not be a simple answer to this technical challenge, but complementary answers involving both the well established 2-DE strategy, and chromatographic methods with improved characteristics (miniaturisation, improved selectivity, stability, ...). For instance, several approaches are needed to explore the pattern of interactions between the components of a given proteome (the 'interactome') [50], including classical chromatography techniques (affinity chromatography, immunopurification combined or not with cross-linking, ...), and proteomics on a chip developments [51–53]. As proteomics is more complex by several orders of magnitude than genomics, no one laboratory or consortium is able to win this challenge alone. In conclusion, future proteomics projects are increasingly likely to result from collaborations between academic and private organisations, built around complementary knowledge and sharing of experimental data and results.

Acknowledgements

N. Imam-Sghiouar is a recipient of a fellowship from Foundation Midy, and F. Poirier is a recipient of a fellowship from the Association pour la Recherche contre le Cancer (ARC).

References

[1] L. Anderson, J.A. Seilhamer, Electrophoresis 18 (1997) 533.
[2] E.S. Lander et al., Nature 409 (2001) 860.
[3] J.E. Celis, P. Gromov, Curr. Opin. Biotech. 10 (1999) 16.
[4] A. Otto, E.C. Muller, E. Brockstedt, M. Schumann, A. Rickers, K. Bommert, B. Wittmann-Liebold, J. Protein Chem. 17 (1998) 564.
[5] R.D. Appel, D.F. Hochstrasser, C. Roch, M. Funk, A.F. Muller, C. Pellegrini, Electrophoresis 9 (1988) 136.
[6] J.I. Garrels, J.T. Farrar, C.B. Burwell, in: J.E. Celis, R. Bravo (Eds.), Two-dimensional Electrophoresis of Proteins, Academic Press, Orlando, 1984, p. 37.
[7] R. Joubert-Caron, J.P. LeCaer, F. Montandon, F. Poirier, N. Imam, J. Feuillard, D. Bladier, J. Rossier, M. Caron, Electrophoresis 21 (2000) 2566.
[8] J. Wojcik, V. Schachter, Brief Bioinform. 1 (2000) 250.
[9] 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, D.F. Hochstrasser, Biotechnology 14 (1996) 61.
[10] P. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[11] D. Figeys, S.P. Gygi, Y. Zhang, J. Watts, M. Gu, R. Aebersold, Electrophoresis 19 (1998) 1811.
[12] P. James, M. Quadroni, E. Carafoli, G. Gonnet, Biochem. Biophys. Res. Commun. 195 (1993) 58.
[13] S.M. Hanash, J.R. Strahler, Y. Chan, R. Kuick, D. Teichroew, J.V. Neel, N. Hailat, D.R. Keim, J. Gratiot-Deans, D. Ungar et al., Proc. Natl. Acad. Sci. USA 90 (1993) 3314.

[14] J.R. Strahler, N. Hailat, B.J. Lamb, K.P. Rogers, J.A. Underhill, R.F. Melhem, D.R. Keim, X. Zhu, R.D. Kuick, D.A. Fox et al., J. Immunol. 149 (1992) 1191.
[15] Y.K. Wang, P.C. Liao, J. Allison, D.A. Gage, P.C. Andrews, D.M. Lubman, S.M. Hanash, J.R. Strahler, J. Biol. Chem. 268 (1993) 14269.
[16] S.M. Hanash, D. Teichroew, Electrophoresis 19 (1998) 2004.
[17] E. Brockstedt, A. Rickers, S. Kostka, A. Laubersheimer, B. Dorken, B. Wittmann-Liebold, K. Bommert, A. Otto, J. Biol. Chem. 273 (1998) 28057.
[18] A. Rickers, E. Brockstedt, M.Y. Mapara, A. Otto, B. Dorken, K. Bommert, Eur. J. Immunol. 28 (1998) 296.
[19] B. Thiede, F. Siejak, C. Dimmler, P.R. Jungblut, T. Rudel, Electrophoresis 21 (2000) 2713.
[20] F. Poirier, N. Imam, M. Pontet, R. Joubert-Caron, R. Caron, J. Chromatogr. B 753 (2001) 23.
[21] D. Lutomski, M. Caron, J.D. Cornillot, P. Bourin, C. Dupuy, M. Pontet, D. Bladier, R. Joubert-Caron, Electrophoresis 17 (1996) 600.
[22] R. Joubert-Caron, J. Feuillard, S. Kohanna, F. Poirier, J.P. LeCaer, M. Schuhmacher, G.W. Bomkamm, A. Polack, M. Caron, D. Bladier, M. Raphael, Electrophoresis 20 (1999) 1017.
[23] F. Poirier, M. Pontet, V. Labas, J.P. LeCaer, N. Imam-Sghiouar, M. Raphael, M. Caron, R. Joubert-Caron, Electrophoresis 22 (2001) 1867.
[24] F. Poirier, M. Pontet, V. Labas, J.P. LeCaer, M. Caron, R. Joubert-Caron, in: Righetti, P.G. (Ed.), ICES 2001: An Electrophoretic Odyssey, 2001, pp. 62. Venona, Italie.
[25] R. Aebersold, J. Leavitt, Electrophoresis 11 (1990) 517.
[26] R.D. Appel, J.R. Vargas, P.M. Palagi, D. Walther, D.F. Hochstrasser, Electrophoresis 18 (1997) 2735.
[27] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850.
[28] http://www.expasy.ch/tools/peptidident.html
[29] W. Zhang, B.T. Chait, Anal. Chem. 72 (2000) 2482.
[30] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551.
[31] V. Santoni, T. Rabilloud, P. Doumas, D. Rouquie, M. Mansion, S. Kieffer, J. Garin, M. Rossignol, Electrophoresis 20 (1999) 705.
[32] S.J. Cordwell, A.S. Nouwens, N.M. Verrills, D.J. Basseal, B.J. Walsh, Electrophoresis 21 (2000) 1094.
[33] P.G. Righetti, A. Castagna, B. Herbert, Anal. Chem. 73 (2001) 320A.
[34] M. Fouillit, F. Poirier, E. Monostori, M. Raphael, D. Bladier, R. Joubert-Caron, M. Caron, Electrophoresis 21 (2000) 275.
[35] M.F. Lopez, B.S. Kristal, E. Chernokalskaya, A. Lazarev, A.I. Shestopalov, A. Bogdanova, M. Robinson, Electrophoresis 21 (2000) 3427.
[36] Y. Oda, T. Nagasu, B.T. Chait, Nat. Biotechnol. 19 (2001) 379.
[37] N.G. Ahn, K.A. Resing, Nat. Biotechnol. 19 (2001) 317.
[38] M. Geng, X. Zhang, M. Bina, F. Regnier, J. Chromatogr. B. Biomed. Sci. Appl. 752 (2001) 293.
[39] S.F. Greer, L.B. Justement, J. Immunol. 162 (1999) 5278.
[40] L.N. Johnson, M. O'Reilly, Curr. Opin. Struct. Biol. 6 (1996) 762.
[41] M. Fouillit, R. Joubert-Caron, F. Poirier, P. Bourin, E. Monostori, M. Levi-Strauss, M. Raphael, D. Bladier, M. Caron, Glycobiology 10 (2000) 413.
[42] L.B. Justement, Adv. Immunol. 66 (1997) 1.
[43] M.R. Larsen, G.L. Sorensen, S.J. Fey, P.M. Larsen, P. Roepstorff, Proteomics 1 (2001) 223.
[44] H. Kaufmann, J.E. Bailey, M. Fussenegger, Proteomics 1 (2001) 194.
[45] A. Sickmann, K. Marcus, H. Schafer, E. Butt-Dorje, S. Lehr, A. Herkner, S. Suer, I. Bahr, H.E. Meyer, Electrophoresis 22 (2001) 1669.
[46] M.R. Wilkins, E. Gasteiger, A.A. Gooley, B.R. Herbert, M.P. Molloy, P.A. Binz, K. Ou, J.C. Sanchez, A. Bairoch, K.L. Williams, D.F. Hochstrasser, J. Mol. Biol. 289 (1999) 645.
[47] A. Bairoch, P. Bucher, K. Hofmann, Nucleic. Acids Res. 25 (1997) 217.
[48] P. Bork, T. Dandekar, Y. Diaz-Lazcoz, F. Eisenhaber, M. Huynen, Y. Yuan, J. Mol. Biol. 283 (1998) 707.
[49] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[50] C. Sanchez, C. Lachaize, F. Janody, B. Bellon, L. Roder, J. Euzenat, F. Rechenmann, B. Jacq, Nucleic Acids Res. 27 (1999) 89.
[51] R.W. Nelson, D. Nedelkov, K.A. Tubbs, Electrophoresis 21 (2000) 1155.
[52] G. MacBeath, S.L. Schreiber, Science 289 (2000) 1760.
[53] D. Figeys, D. Pinto, Electrophoresis 22 (2001) 208.