Journal of Pharmaceutical and Biomedical Analysis 53 (2010) 50–57

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Characterization of factor VIII pharmaceutical preparations by means of MudPIT proteomic approach

Fabrizio Basilico[a,1], Ilaria Nardini[b,1], Filippo Mori[b], Elena Brambilla[a], Louise Benazzi[a], Antonella De Palma[a], Enrico Rosti[a], Claudio Farina[b], PierLuigi Mauri[a,*]

[a] ITB/CNR, via Fratelli Cervi 93, 20090 Segrate (MI), Italy
[b] Kedrion S.p.A, via Provinciale 1, 55027 Gallicano (LU), Italy

ARTICLE INFO

Article history:
Received 11 December 2009
Received in revised form 19 March 2010
Accepted 22 March 2010
Available online 27 March 2010

Keywords:
Coagulation factor VIII
Protein profiles
MudPIT
Two-dimensional chromatography
Mass spectrometry

ABSTRACT

For a good clinical outcome of Haemophilia A substitutive therapy a detailed characterization of factor VIII (FVIII) concentrates is required, in order to disclose the eventual relations between differently composed concentrates and their biological effects. This preliminary work could be a first step towards a deep structural characterization of FVIII concentrates, using the fast and simply manageable MudPIT technology, which enables the identification and characterization of protein mixtures taking advantage of both the high separation capacity of two-dimensional chromatography and the powerful peptide characterization ability of tandem mass spectrometry. The aim of this study was to evaluate the suitability of for the characterization of FVIII molecule in complex mixtures such its commercial concentrates, both plasma-derived and recombinant, and for the determination of the protein composition of different FVIII preparations. By means of Multidimensional Protein Identification Technology (MudPIT) it was possible to assess the presence of factor VIII in its preparations and to identify most of the contaminant proteins without gel separation.

In particular, 125 and 42 proteins were identified in plasma-derived and recombinant concentrates, respectively. Concerning investigation of FVIII, 24 different peptides were identified in plasma-derived corresponding to 7, 29, 27, 19 and 67 of percentage coverage for A1, A2, A3, C1 and C2 domains, respectively. About its multimeric carrier von Willebrand factor (VWF), we have sequenced 42% of domain interacting with A3 and C2 domains of FVIII. Finally, it has been observed that normalized parameters, such as total peptide hits obtained by SEQUEST may be used for evaluation of the relative abundance of FVIII in different preparations.

1. Introduction

Haemophilia A is the most common severe life-long bleeding disorder. It is caused by an inherited deficiency of factor VIII (FVIII) due to mutations in the corresponding gene. Its severity correlates with the residual FVIII activity found in plasma and depends on the type of FVIII gene mutations [1]. In its most serious forms, when FVIII activity is below 1% of normal, this disorder is life-threatening [2]. In such affected individuals, haemostasis may be restored by replacing the missing or defective protein with FVIII either isolated from natural source or expressed as a recombinant analogue in mammalian cells [3].

FVIII is a large multi-domain glycoprotein with domain structure A1-α1-A2-α2-B-α3-A3-C1-C2, and is involved in the cascade of biochemical reactions leading to coagulation [4,5]. In the serum, FVIII is cleaved into heavy (A1-α1-A2-α2-B) and light chains (α3-A3-C1-C2), which are held together by a divalent metal ion [4,6]. This heterodimer circulates in association with von Willebrand factor (VWF), a large multimeric glycoprotein whose monomers are held together by disulfide bridges and non-covalent interactions [7,8]. The stoichiometry of factor VIII and VWF in plasma is approximately 1:50; therefore, VWF represents most of factor VIII–VWF complex in terms of weight [9]. Binding of factor VIII to VWF is essential for the survival, stabilization and function of FVIII in vivo [5]. The underlying mechanism probably relies on the fact that FVIII is bound to VWF and is therefore protected from phospholipid-dependent proteolysis by activated protein C and factor Xa [10,11].

Because FVIII and VWF form a tightly bound non-covalent complex in plasma [12], both proteins are co-purified when isolated from plasma, unless special measures are taken [13]. FVIII cir-

* Corresponding author. Tel.: +39 0226422728; fax: +39 0226422770.
E-mail address: pierluigi.mauri@itb.cnr.it (P. Mauri).
[1] These authors equally contributed to this work.

0731-7085/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2010.03.027

culates in plasma at a very low concentration (0.1–0.2 μg/ml), and represents only a small proportion of plasma protein content. In plasma-derived preparations (pdFVIII), often defined as high-purity products, FVIII still represents only 1–2% of total protein content. PdFVIII concentrates are in fact complex protein mixtures, as they contain other plasma proteins such as fibrinogen and fibronectin [11–14].

Recombinant preparations (rFVIII) contain instead FVIII highly purified, but they require the addition of large amounts of stabilizers such as human serum albumin (HSA) or VWF [15].

One of the most important complications in patients treated with FVIII concentrates is the development of antibodies inhibiting FVIII activity. The latter are called inhibitors, and are directed towards the A2, C2 and A3 domains of the molecule. About 25–30% of patients develop antibodies inhibiting FVIII activity. The inhibitory effects of antibodies occur at various stages of the FVIII functional pathway, including FVIII binding to VWF, activation of FVIII by thrombin, and activated FVIII (FVIIIa) incorporation into the Xase complex [16].

The detailed analysis of FVIII preparation composition and of the structure of FVIII itself is therefore crucial in order to improve the clinical outcome of FVIII replacement therapy. Because FVIII preparations are complex protein mixtures, their analysis has represented a challenge, due to the relatively low levels of this protein in concentrates, the high abundance of VWF and the presence of other plasma proteins [13,14].

Clinical chemistry assays of FVIII preparations are based on the clotting and chromogenic methods, which determine the FVIII procoagulant activity [17], and ELISA methods and immunoblotting which allow the determination of FVIII:Ag. Molecular characterization of FVIII preparations is based on electrophoresis (mono-, 1D and bi-dimensional, 2D) [18], chromatography (like size exclusion, immunoaffinity and reversed phase) [19] and mass spectrometry using MALDI-TOF [18] or electrospray usually coupled to liquid chromatography (LC/ESI-MS) [19].

However, these methodologies give a poor characterization of the protein content of FVIII enriched preparations. A new proteomic methodology based on two-dimensional capillary chromatography coupled to tandem mass spectrometry (2DC-MS/MS), also named Multidimensional Protein Identification Technology (MudPIT), has been developed recently [20]. This approach does not require gel separation and it is a mass spectrometry-based method. It involves the generation of peptides by enzymatic digestion of a complex protein mixture, their separation by means of two micro-HPLC columns (cation exchange and reversed phase, 2DC) and direct analysis of eluted peptides by tandem mass spectrometry (MS/MS). Using an appropriate software, such as the SEQUEST algorithm, based on sequence database searching, the full and MS/MS spectra are then correlated to specific peptide sequences and the corresponding protein are identified [21]. MudPIT takes advantage of both the high separation capacity of 2DC (enabling the characterization of proteins with extreme isoelectric point, molecular weight or hydrophobicity) and peptide identification by amino acid sequencing. In addition, MudPIT technology is reported to supply quantitative data [22,23].

The aim of this study was to evaluate the suitability of a MudPIT approach for the characterization of FVIII in its commercial preparations, and determination of the protein profiles of these complex mixtures. An in-depth and rapid characterization of FVIII concentrates could disclose an eventual relation between different batches or products and their biological effects.

Besides the identification of an adequate coverage of FVIII sequence, we were able to identify most of the contaminant proteins, including those unlikely to be detected with other traditional and labour-intensive approach like the electrophoretic one.

2. Materials and methods

2.1. FVIII concentrates

Emoclot D.I. (Kedrion S.p.A., Lucca, Italy) is an high-purity, plasma-derived FVIII concentrate (pdFVIII) obtained through anion-exchange chromatography (1000 IU, 10 mg total protein content); a commercial recombinant FVIII (rFVIII; 1000 IU) (Baxter International Inc., Deerfield, IL, USA) obtained from a genetically engineered Chinese Hamster Ovary cell line, stabilized with human serum albumin (HSA 12.5 mg/ml), was used [24].

2.2. Samples preparation

Due to its high amount of albumin, the rFVIII sample was treated with Microcon YM-100 (cut-off 100 kDa) Centrifugal Filter Device (Millipore, Billerica, MA, USA) in order to separate FVIII from HSA, whose high content hampers FVIII detection. After the centrifugation (14,000 × *g* for 12 min) the filter was subjected to 4 washing steps with ammonium bicarbonate 100 mM pH 8.

Treated recombinant and plasma-derived samples were diluted (1:1) with RapiGest™ SF 1% (Waters, Milford, MA, USA). Proteins/RapiGest™ SF mixtures (both rFVIII and pdFVIII) were heated to ∼100 °C for 5 min and then samples were collected before reduction and alkylation. PdFVIII and rFVIII samples (50 μL) were reduced with 20 mM dithiothreitol (DTT, Sigma–Aldrich, Milan, Italy), at 37 °C for 4 h. The samples were then alkylated by the addition of iodoacetamide (IAA, Sigma–Aldrich, Milan, Italy) to 100 mM and incubated at the same temperature, in the dark, for approximately 1 h.

Sequencing grade modified trypsin (Promega, Madison, WI, USA) was added at enzyme–substrate ratio around 1:50 (w:w). After overnight incubation at 37 °C, the reaction was stopped by acidification with trifluoroacetic acid (Sigma, St. Louis, MO, USA) to a final concentration of 0.5%, pH ∼1.5. Trifluoroacetic acid also causes RapiGest™ SF inactivation. Samples were incubated at 37 °C for 45 min and then centrifuged at 14,000 × *g* for 10 min. Peptide mixtures were purified and concentrated using PepClean™ C-18 Spin Columns (Pierce, Rockford, IL, USA). Samples were dried in Speed-Vac at 60 °C. 20 μl of 5% ACN/0.1% formic acid solution was then added.

2.3. 2DC-MS/MS analysis

Digested pdFVIII and rFVIII samples were analyzed by means of two-dimensional microchromatography coupled to ion trap mass spectrometry (2DC-MS/MS, also referred to as MudPIT), Briefly, 5 μl of the digested peptide mixture was first separated by ion exchange chromatography (BioBasic-SCX column, 0.32 i.d. × 100 mm, 5 (m, Thermo Electron Corp., Bellofonte, PA, USA) by applying a ten-step ammonium chloride concentration gradient (5, 10, 15, 20, 30, 40, 80, 120, 400, 700 mM). Each salt step eluate was directly loaded on a C_{18} reversed-phase column (BioBasic-18, 0.180 i.d. × 100 mm, 5 (m, Thermo Electron Corp., Bellefonte, PA, USA) and separated with an acetonitrile gradient (eluent A, 0.1% formic acid in water; eluent B, 0.1% formic acid in acetonitrile); Each of the 10 cycles had the following profile: 5 min of 5% B, a 45 min gradient from 5 to 65% B, 3 min of 65% B, a 10 min gradient from 65 to 95% B, 5 min of 95% B, a 5 min gradient from 95 to 5% B, and finally 17 min of re-equilibration with 5% B; flow rate, 2 μL/min.

Peptides eluted from the C_{18} column were analyzed directly with a linear ion trap LTQ mass spectrometer equipped with nano-ESI source. The heated capillary was held at 185 °C, ion spray 1.5 kV and capillary voltage 47 V. Spectra were acquired in positive ion mode (in the range of 400–2000 m/z) using dynamic exclusion for MS/MS analysis (collision energy 35%). The mass spectra were

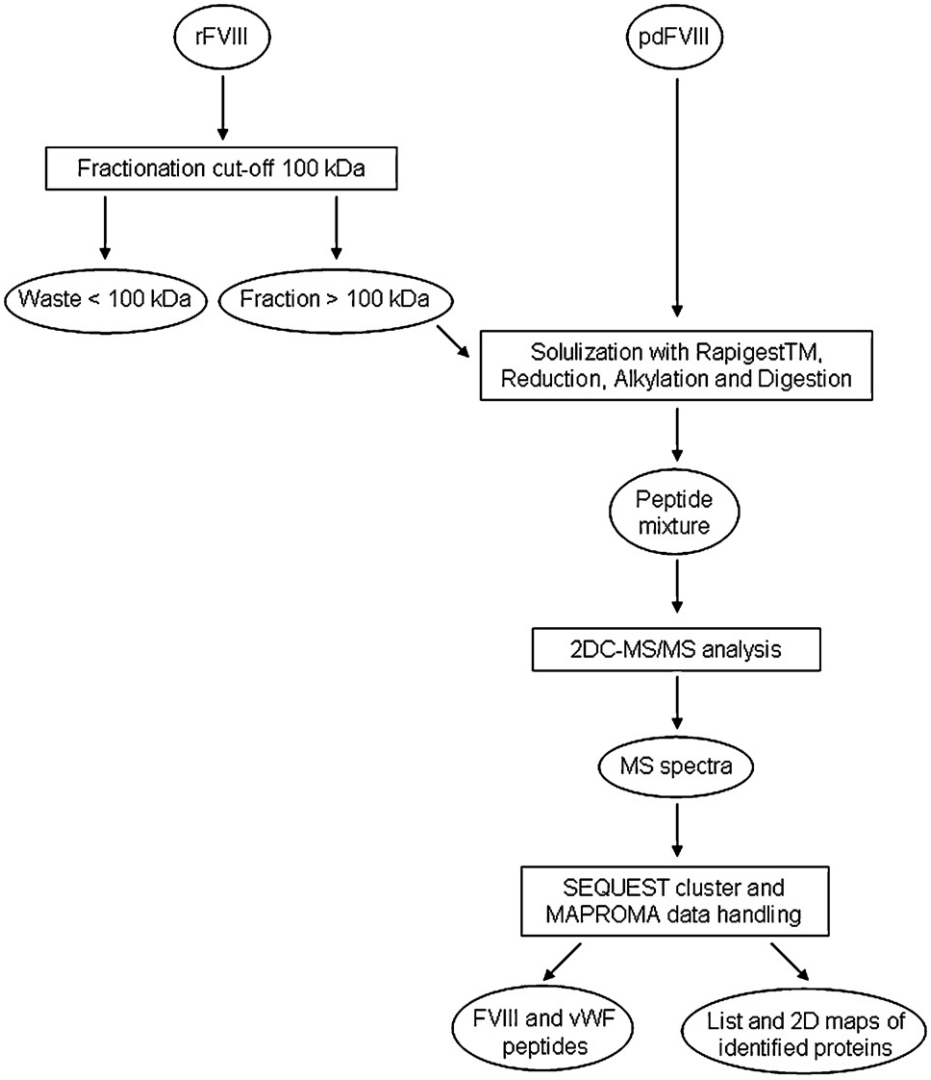

Fig. 1. Main steps used for determining the protein profiles of rFVIII and pdFVIII pharmaceutical preparations.

acquired in data dependent mode (full MS scan followed by five MS/MS scans on the five most intense ions).

2.4. Mass spectrometry data handling

The experimental mass spectra (full MS and MS/MS) were correlated to peptide sequences by comparison with the theoretical mass spectra deduced from human (193,189 entries) and hamster (221 entries) protein databases obtained as subsets of nr (non-redundant) database from the NCBI website (www.ncbi.nlm.nih.gov/Ftp/index.html), and downloaded on March 2008. Computer analysis of MS/MS spectra was performed using the Bioworks software, version 3.3.1, based on SEQUEST cluster as search engine (University of Washington, USA, licensed to Thermo Electron Corp., San Jose, CA, USA). For peptide matching the following limits were used: Xcorr scores greater than 1.5 for singly charged peptide ions and 2.0 and 2.5 for doubly and triply charged ions, respectively; peptide mass search tolerance was set to 1.0 Da.

To assign final score to proteins, the SEQUEST output data were filtered using the following parameters: peptide probability to 10^{-3} and consensus score higher than 10. False positive was estimated by means of reverse database and the percentage ratio between reverse and native database matches was less than 2.5%.

The output data obtained from SEQUEST software were treated with the MAProMA (Multidimensional Algorithm Protein Map) in-house algorithm for comparison of protein lists, evaluation of relative abundances, and plotting of virtual 2D maps [25].

3. Results

The main analytical steps applied in the present work are summarized in Fig. 1. Rapigest detergent, compatible with mass spectrometry analysis, was used for protein dissolution and denaturation. After reduction and alkylation each sample was digested overnight with trypsin, and the resulting peptide mixture was analyzed by means of MudPIT system, which allows a two-dimensional chromatographic separation and direct loading into ion trap mass spectrometer. Because rFVIII samples contained a high amount of albumin, added as stabilizer (higher than 99.9%), a pre-treatment based on ultrafiltration (cut-off 100 kDa) was introduced to reduce albumin.

3.1. Analysis of pdFVIII preparation

Direct digestion and MudPIT analysis, without detergent application, permitted the characterization of few proteins present in pdFVIII samples, and only one FVIII unique peptide was identi-

Table 1
List of most frequently (≥4) identified proteins in plasma-derived FVIII preparations.

NCBI accession	Reference	pI	MW	Score[a]	Hits[a,b]	Frequency[c]
401413	von Willebrand factor (VWF)	5.2	309,299	534	53.4	6
2506872	Fibronectin precursor (FN)	5.4	262,606	467	46.7	6
4503647	Coagulation factor VIII	6.9	267,009	155	15.5	6
125000	Inter-alpha-trypsin inhibitor heavy chain H2	6.4	106,437	137	13.7	6
2851501	Inter-alpha-trypsin inhibitor heavy chain H1	6.3	101,389	125	12.5	6
2781209	Chain C, crystal structure of fibrinogen	5.8	36,181	120	12.0	6
4503689	Fibrinogen, alpha polypeptide	5.6	94,973	92	9.0	6
87919	Ig mu chain precursor	5.8	68,511	74	7.3	6
24987624	1LT9_B	7.0	35,899	64	6.3	6
3024064	Inter-alpha-trypsin inhibitor heavy chain H3	5.5	99,122	55	5.5	6
5031863	Galectin 3 binding protein	5.0	65,332	52	5.2	6
42740907	Clusterin isoform 2	5.9	52,495	52	5.2	6
18655500	1JPS_L	7.0	23,500	42	4.2	6
4389275	Human serum albumin	5.6	66,036	40	4.0	6
114014	Apolipoprotein B-100	6.6	515,560	35	3.5	6
47419932	Platelet glycoprotein Ib alpha	6.2	68,956	34	3.3	6
88853069	Vitronectin precursor	5.5	54,306	32	3.2	6
81175238	Complement C4-B precursor	6.6	192,771	30	3.0	6
4699843	1BIK	4.7	15,975	30	3.0	6
70906435	Fibrinogen, beta chain	8.2	55,929	27	2.7	6
2146957	Ig heavy chain V region precursor	7.4	41,784	24	2.3	6
78101267	Chain A, human complement component C3	6.9	71,191	22	2.2	6
106659	Ig lambda chain	8.0	24,834	19	1.8	6
3660204	Chain A, structure of recombinant alphaec domain	4.2	22,825	19	1.8	6
50659080	Serpin peptidase inhibitor	5.2	47,652	19	1.8	6
70906439	Fibrinogen, gamma chain	5.3	51,512	17	1.7	6
87191	B20807 complement C4-B	5.7	19,345	12	1.2	6
139641	Vitamin D-binding protein	5.3	52,964	12	1.2	6
2781208	Chain B, crystal structure of fibrinogen	5.8	37,650	10	1.0	6
11514600	1E8B_A	5.9	17,808	10	1.0	6
31615935	1OW0_A	7.1	23,201	16	1.6	5
24987623	1LT9_A	8.9	7734	16	1.6	5
50363217	Serine (or cysteine) proteinase inhibitor	5.3	46,737	14	1.4	5
125507	Kininogen-1	6.3	71,946	23	2.3	4
4505047	Lumican precursor	6.2	38,430	15	1.5	4
48429255	Extracellular matrix protein 1 precursor	6.2	60,675	15	1.5	4
7428712	Keratin 1, type II	6.0	65,494	15	1.3	4
4557287	Angiotensinogen preproprotein	5.9	53,155	13	1.3	4
115583663	Alpha-2-plasmin inhibitor	5.8	54,566	13	1.3	4
123858	Ig heavy chain V–III region	10.1	12,432	10	1.0	4
108860890	Plexin-A4 precursor	6.4	21,2455	10	1.0	4

[a] Average values.
[b] Distinct peptides.
[c] Out of 6 samples.

fied by MS/MS sequencing. This poor characterization is probably due to the presence of VWF, which forms a tightly bound complex with FVIII, protecting it from tryptic digestion. For this reason, Rapigest detergent was used for protein complex dissolution and to increase both the number of identified proteins and percentage of sequenced FVIII.

Typically, for each MudPIT analysis of pdFVIII preparation, about 65 proteins were characterized and around 45% of these were identified by two or more unique peptides. Considering all analyzed samples, 73 different proteins out of 125 identified were detected at least twice (see Table I in supplementary data). The identification of some proteins, such as inter-alpha-trypsin inhibitor heavy chains H1, H2 and H3, fibrinogen, vitronectin, coagulation factor II precursor, fibronectin, gp1bα, Ig and plasminogen, confirmed previous characterization carried out by other authors [26]. Other proteins, such as apolipoprotein B, complement C4-A and complement component C3, were never reported in pdFVIII preparations. Table 1 reports the most frequently identified proteins (frequency ≥ 4 out of 6 samples) and their frequency.

Protein list output data have been plotted on a 2D map using MAPROMA software. The latter assigns to each protein a colour/shape code according to the range score values or the number of identified peptides obtained by SEQUEST data analysis [27], or in relation to their frequency. For example, Fig. 2 shows the 2D

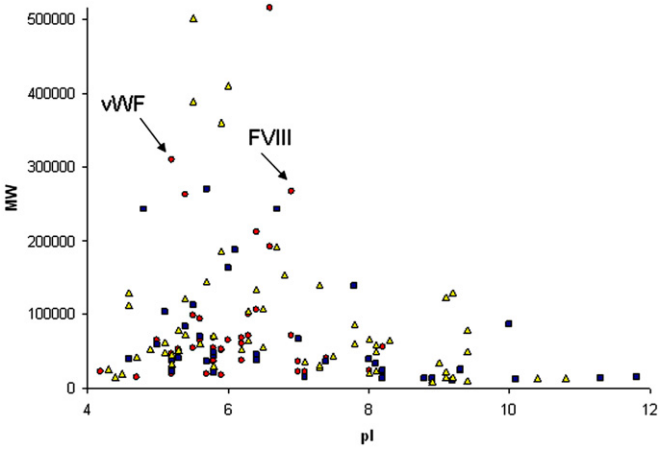

Fig. 2. Two-dimensional map (2D map) obtained by plotting theoretical pI vs MW over six different MudPIT analyses of pdFVIII (see Table I of supplementary data). Colour/shape codes of spots are related to frequency of identified proteins: yellow/triangle, blue/square and red/circle correspond to 1, from 2 to 3, and ≥4 times out of 6, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

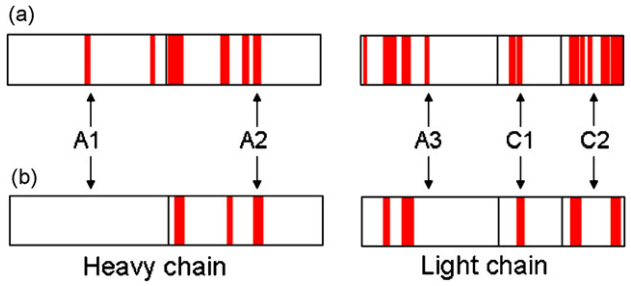

Fig. 3. Sequence coverage of FVIII protein obtained by MudPIT analysis of (a) plasma-derived and (b) biotechnological preparations.

map obtained by plotting (theoretical pI vs MW) the protein list reported in Table I of supplementary data. In this way it is possible to obtain an overview of protein distribution and a rapid evaluation of detection frequency for each protein out of 6 different analyses. Also, the 2D map reported in Fig. 2 underlines that the MudPIT approach allows the identification of alkaline proteins, such as histones (pI around 11–11.8), or high molecular weight proteins, such as apolipoprotein B-100 (around 500 kDa) and serine/threonine-protein kinase SMG1 (around 400 kDa). Concerning the FVIII molecule, the MudPIT analysis allowed the identification of 24 different peptides of FVIII, corresponding to about 18% of heavy chain (7% A1, 29% A2 domains) and 27% of light chain (27% A3, 19% C1, 67% C2 domains) sequences (see Fig. 3). Among these, 7 peptides were always detected in the 2DC-MS/MS analyses (see Table 2).

Concerning VWF, the physiological carrier of FVIII, 57 different peptides were identified (see supplementary data, Table II), corresponding to about 25% of the entire protein sequence. In particular, about 42% of the VWF domain interacting with A3 and C2 domains of FVIII has been sequenced (see Table 3).

Recent results indicate that mass spectrometry-based methodologies have had a significant impact on the ability to apply quantitative analysis to proteomic studies using the so-called "label-free" approach. In particular, it was reported that peptide count [28], percentage coverage [23] or score value [27] allow relative evaluation of protein content. In this context, we compare the total identified peptides for FVIII and VWF in one batch and its dilution (1:2), with the peak area of the extracted ion chromatograms (EIC) for selected peptides of each protein. Specifically, for each protein 5 peptides always detected were chosen. The ratio of total peptides for FVIII (2.0) and VWF (1.8) resulted in good agreement with the ratios of extracted areas (2.2 and 1.9, for FVIII and VWF respectively). Based on these findings we used VWF protein as a reference for evaluating the experimental variation in FVIII detection. In particular, for each MudPIT analysis the VWF/FVIII ratio was calculated using different parameters provided in the output of SEQUEST data handling. More specifically, VWF/FVIII ratios were evaluated in terms of total identified peptides (17.3), protein score values (3.4) and sequence coverages (9.1); replicate analyses ($n = 6$) showed good reproducibility (coefficient of variation, C.V.%, less than 12%). Finally, considering score values, normalized using VWF as internal reference, the analyses of batch replicates (Fig. 4) allowed a good reproducibility (slope ~ 0.99) and regression coefficient ($r^2 \sim 0.98$).

3.2. Analysis of recombinant rFVIII preparation

The biotechnological FVIII preparations under consideration contained human serum albumin (HSA), added as stabilizer of recombinant FVIII protein; as a consequence, direct MudPIT analysis of digested rFVIII concentrate allowed the detection of few proteins (almost exclusively albumin) and only one peptide related to FVIII was identified. For this reason, it was necessary to reduce the amount of HSA in the rFVIII sample. Since the employment of albumin depletion kits caused even a relevant loss of FVIII (data not shown), the reduction of albumin content was achieved by means of a centrifugal filter device with a cut-off around 100 kDa. In particular, the sample was fractionated in two parts: the filtered fraction, containing about 95% of total HSA, and the retained fraction, containing the FVIII protein (see Fig. 1). MudPIT analysis of digested recovered rFVIII fraction and data handling based on human protein database allowed the identification of 42 proteins

Table 2
Identified peptides of FVIII protein ($n = 6$ MudPIT analyses) in pdFVIII preparation; peptides identified in all analyses are in bold. C* corresponds to alkylated cysteine.

Domains	Peptide location	Peptide sequence	Frequency in pdFVIII samples
A1 domain	**186–199**	**DLNSGLIGALLVC*R**	**6/6**
	345–355	VDSC*PEEPQLR	3/6
A2 domain	379–391	FDDDNSPSFIQIR	4/6
	400–424	WVHYIAAEEEDWDYAPLVLAPDDR	4/6
	519–529	**DFPILPGEIFK**	**6/6**
	532–542	WTVTVEDGPTK	4/6
	561–576	**DLASGLIGPLLIC*YK**	**6/6**
	591–602[a]	NVILFSVFDENR	5/6
	603–612[a]	SWYLTENIQR	4/6
B domain	1032–1059	HIDGPSLLIENSPSVWQNILESDEFK	1/6
A3 domain	1716–1724	HYFIAAVER	3/6
	1752–1768[a]	VVFQEFTDGSFTQPLYR	4/6
	1769–1783	GELNEHLGLLGPYIR	1/6
	1796–1816[a]	NQASRPYSFYSSLISYEEDQR	2/6
	1801–1816	PYSFYSSLISYEEDQR	1/6
	1853–1864	**AWAYFSDVDLEK**	**6/6**
C1 domain	**2053–2068**[a]	**DFQITASGQYGQWAPK**	**6/6**
	2072–2084	**LHYSGSINAWSK**	**6/6**
C2 domain	2203–2226	AISDAQIASSYFNMFAWSPSK	1/6
	2235–2246	SNAWRPQVNNPK	2/6
	2259–2268	VGVQGVK	2/6
	2278–2298	EFLISSSQDGHQWLFFQNGK	3/6
	2301–2323[a]	VFQGNQDSFTPVVNSLDPPLLTR	2/6
	2327–2339	**IHPQSWVHQIALR**	**6/6**

[a] Peptides identified in rFVIII sample, also.

Table 3
Identified peptides of VWF protein involved in the interaction with FVIII A3 and C2 domains (frequency is related to $n = 6$ MudPIT analyses). C* corresponds to alkylated cysteine.

VWF peptide location	Identified peptide sequence	Frequency in pdFVIII samples
774–782	LVC*PADNLR	3/6
844–852	IGC*NTC*VC*R	3/6
883–906	YLFPGEC*QYVLVQDYC*GSNPGTFR	5/6
925–944	VTILVEGGEIELFDGEVNVK	4/6
961–968	YIILLLGK	2/6
977–985	HLSISVVLK	5/6
992–1026	VC*GLC*GNFDGIQNNDLTSSNLQVEEDPVDFGNSWK	2/6

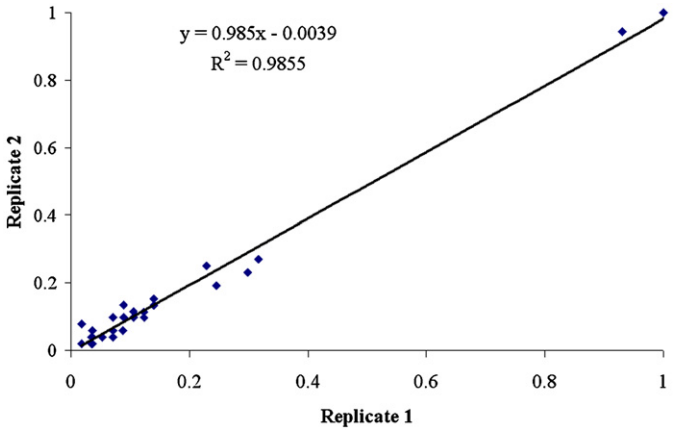

Fig. 4. Plotting of protein score ratio (using VWF as internal standard and protein identified with 2 or more distinct peptides) from two replicates $n = 43$.

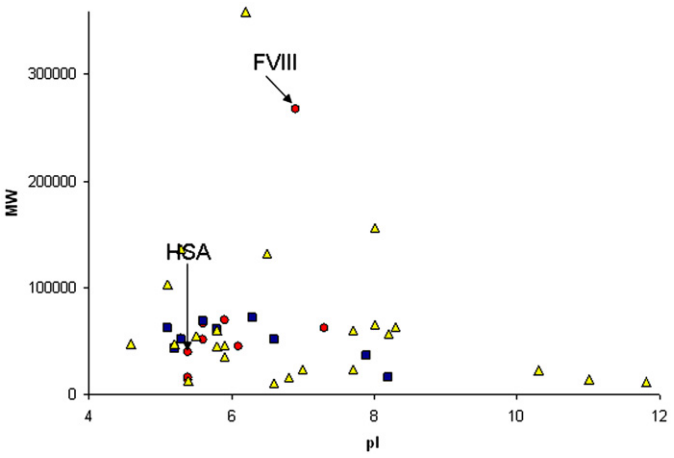

Fig. 5. 2D map of identified protein over six different MudPIT analyses of rFVIII (see Table III of supplementary data). Colour/shape codes of spots are related to frequency of identified proteins: yellow/triangle, blue/square and red/circle correspond to 1, from 2 to 3, and ≥4 times out of six, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

(see Table III in supplementary data), and 9 of these (including albumin and FVIII) were identified in each analysis; 14 proteins were identified at least in 2 analyses out of 3 (in duplicate injection).

Because rFVIII preparations contain HSA as stabilizer, whereas the cell line used for FVIII synthesis was CHO (Chinese Hamster Ovary), both the database, human and hamster, were used for SEQUEST data handling. In particular, no proteins were identified from the hamster database, whereas all the proteins identified (shown in Table III in supplementary data) were typical of human serum. These results indicate that all contaminant proteins identified derived from added HSA and not from hamster cells used for biotechnological production of FVIII. The proteins identified in rFVIII sample were plotted on a 2D map (theoretical pI vs MW) in relation to their frequency (see Fig. 5).

Concerning the FVIII protein, about 8% of its sequence was characterized. Specifically, 11 distinct peptides related to FVIII were identified (see Table 4 and Fig. 3B) belonging to A2, A3, C1 and C2 domains. Even in the case of the recombinant product the largest sequence coverage was obtained for the C2 domain (≈30%). One identified peptide was detected in each analysis and 5 peptides were also identified in the analyses of pdFVIII concentrates (compare Tables 2 and 4).

4. Discussion

FVIII pharmaceutical preparations, both plasma-derived (pdFVIII) or produced by DNA recombinant technology (rFVIII), are complex protein mixtures: in plasma-derived concentrates, FVIII represent only 1–2% of total protein content, whereas recombinant preparations may contain other proteins as stabilizers. In addition, protein composition of FVIII concentrates depends both on the source used (e.g. human plasma) and the degree of purification achieved by means of the different production procedures. In this

Table 4
Identified peptides of FVIII protein in rFVIII ($n = 6$ MudPIT analyses); peptide identified in each analysis is in bold.

Domains	Peptide location	Peptide sequence	Frequency
A2 domain	400–424	TWVHYIAAEEEDWDYAPLVLAPDDR	4
	530–542	YKWTVTVEDGPTK	1
	591–602[a]	NVILFSVFDENR	3
	591–613	NVILFSVFDENRSWYLTENIQR	4
	603–612[a]	SWYLTENIQR	2
A3 domain	1752–1768[a]	VVFQEFTDGSFTQPLYR	2
	1796–1816[a]	NQASRPYSFYSSLISYEEDQR	1
	1796–1824	**NQASRPYSFYSSLISYEEDQRQGAEPR**	6
C1 domain	2053–2068[a]	DFQITASGQYGQWAPK	4
C2 domian	2203–2228	AISDAQITASSYFTNMFATWSPSKAR	5
	2301–2323[a]	VFQGNQDSFTPVVNSLDPPLLTR	1

[a] Peptides identified in pdFVIII sample, also.

Fig. 6. 3D structure of A1 (fuchsia), A2 (violet), A3 (green), C1 (light-blue) and C2 (pink) domains. Identified peptides are reported in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

context, besides the identification of an adequate coverage of FVIII molecule, it is of primary importance to characterize the other proteins occurring in FVIII concentrates.

Previous works used gel-based analysis to investigate the protein composition of FVIII pharmaceutical preparations [14,26]. On the contrary, the analytical strategy employed in the present investigation used 2DC-MS/MS (also called MudPIT) proteomic approach, a mass spectrometry-based method, both to characterize the FVIII protein and to determine the protein profiles of plasma-derived and recombinant concentrates of FVIII.

Our results indicate that the MudPIT approach allows the characterization of a good coverage of FVIII sequence in its pharmaceutical preparations. In particular, the peptides of FVIII which were identified are localized mainly on A2, A3 and C2 domains. This is of great interest, because such sites are those which are involved in interactions with VWF and inhibitors. Although the same amount of FVIII protein was contained in pdFVIII and rFVIII concentrate samples used for the analyses, a different number of peptides were detected; specifically, 24 and 11 FVIII distinct peptides were identified in plasma and biotechnological preparations respectively. Only five sequenced peptides of FVIII in biotechnological preparations are shared with those obtained from pdFVIII concentrates. The low number of FVIII peptides identified in biotechnological preparations could be due to the presence of HSA, which, although reduced, may compete with FVIII for tryptic digestion to a greater extent than VWF. Another explanation of the poor sequencing of FVIII may lie in different post-translational modifications occurring in the recombinant product compared to the plasma-derived one, which could hinders FVIII peptides identification.

Fig. 6 reports the three-dimensional structure of FVIII, except for the B domain. It is derived from the work of Stoilova-McPhie et al. [29]. The identified FVIII peptides obtained by MudPIT analysis (in yellow) are located mainly on the exposed and unstructured domains of the protein. It is worthy of notice that the A3 and C2 domains, which are the main target of antibodies reaction and are responsible for interactions with VWF, were extensively sequenced. The low coverage of B domain could be indeed due to its partial or nearly complete cleavage in FVIII mature heterodimer, besides the presence of several glycosylation sites, which does not allow the sequencing of the involved peptides, if not subjected to deglycosylation protocols. For these reasons, in plasma-derived preparation only 3% of B domain has been detected.

Concerning the protein profiles of FVIII concentrates obtained by the MudPIT approach, it was confirmed that both pharmaceutical preparations we analyzed contain several proteins deriving from plasma source. Specifically, more plasma proteins were identified in pdFVIII concentrates (more than 100 proteins) in comparison with biotechnological ones (about 40 proteins). This is probably due to the relative low purification of FVIII from plasma; on the contrary, in rFVIII preparations plasma proteins are correlated to albumin, purified form plasma source, which was added to the final formulation. Specifically, the MudPIT method confirmed the presence of proteins previously identified by gel-based approach [14], and detected many other plasma ones. In particular, in pdFVIII concentrates many proteins with high MW or extreme pI values were identified (see Fig. 2). Quantitative analysis of proteomes can be obtained by means of specific tags, such as SILAC (stable-isotope labeling with amino acids in cell culture), ICAT (isotope code affinity tag assay) methodologies [30] and iTRAQ (isobaric tag for relative and absolute quantitation) [26] or using label-free approach [31]. In general, stable-isotope labeling has an undisputable benefit as the comparison between two or more samples is performed in the same single analysis. There are, however, some inherent drawbacks to this approach, due to the additional steps and costs required for protein labeling, and its efficiency. In general, the label-free strategy is more straightforward and less expensive. It appears to be particularly useful for direct and high-throughput comparison of samples in a fully automated setting. The main feature of the label-free approach is that the samples of interest are analyzed in distinct replicate analyses.

By analyzing different replicate samples, we observed a reproducible ratio of score values, number of identified peptides and sequence coverage obtained by the SEQUEST algorithm. This suggests that these parameters are related to quantitative changes in the amount of proteins detected in different samples, and confirms previous work on the possibility of obtaining, from MudPIT analysis, a quantitative evaluation of the relative protein abundance, in conjunction with database engine outputs such as spectral sampling [22] and peptide hits [28]. In addition, by normalizing parameter values obtained from SEQUEST, a quantitative evaluation of relative protein abundance is possible. The procedure is relatively simple, can be automated, does not require expensive reagents or tedious preparation steps, and only a small amount of sample is necessary (50 μl). Our findings suggest that it may be possible to use the output SEQUEST data for monitoring the quality of FVIII pharmaceutical production. It should be stressed that the ratio between protein identification parameters (e.g. total peptides) concerns the relative amounts of the different proteins within a sample, but does not provide absolute estimates. Determination of absolute protein content is more difficult and requires to use proteotypic peptides as internal standards, labelled and added into the complex mixture to obtain calibration curves [32].

It is interesting to note that the molar ratio (about 16) between VWF and FVIII, inferred from VWF antigen test and FVIII antigen [33] determination, is similar to peptide ratio resulting by MudPIT analysis (about 17). This agreement between results obtained with different techniques represent a further evidence of MudPIT quantitative potential, which needs more investigation.

The 100 kDa ultrafiltration step for rFVIII concentrates allowed the removal of 95% of HSA and the detection of 11 peptides correlated to FVIII; however, due to low reproducibility of albumin depletion, the evaluation of FVIII/stabilizer (HSA) ratio in the resulting fraction was not performed.

In conclusion, we have demonstrated that the proposed method is a suitable approach for the characterization of protein profiles of FVIII pharmaceutical preparations, and also for the detection of the main immunological domains of FVIII protein. The MudPIT proteomic approach may be used for a rapid screening of different preparations and batches in relation to quality control, determination of the protein components which could be involved

in immunomodulation [34] and possibly the investigation of the correlation between protein profiles and biological activity. The proposed method can also be used to detect protein impurities deriving both from the cell culture and the stabilizer added in rFVIII concentrates. This latter feature is not of minor importance, because recombinant products have been developed to avoid the risk of transmission of infectious diseases from plasma source. However, they may contain HSA and related contaminants deriving from plasma, which are also a potential source of viral infectivity [35] and immunogenic activity.

In addition on the basis of the ratio between total identified peptides it could even be possible to forecast a relative quantification of the main proteins, as suggested by the close agreement between this value and the molar ratio figure between FVIII and VWF inferred by means of their traditional assays.

Concerning FVIII analysis, the MudPIT proteomic approach provides a significant improvement over gel-based approach, because it represents a fully automated technology. The present study represents a proof of principle for future applications of MudPIT proteomic methodology in the determination of protein profiles and in the quality control of protein-based pharmaceutical preparations. Specifically, when applied to FVIII pharmaceutical concentrates, this method may be used to characterize FVIII and impurities in each production batch. This is very important for future screening of protein composition and immunological effects of FVIII preparations.

In the next future, this technology could provide a useful tool for a wide and in-depth characterization of FVIII molecule (e.g. phosphorylation, glycosylation and secondary structure) and of the main contaminant proteins included in its concentrates. This may help us explain the molecular basis for the biological and clinical effects (e.g. the immunogenic response) of replacement therapies carried out with different FVIII-based products.

Acknowledgments

This work was supported (to PLM) by Fondazione CARIPLO (Proteomic platform, Operational Network for Biomedicine Excellence in Lombardy project). The authors thank Marta G. Bitonti for MAProMA software.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jpba.2010.03.027.

References

[1] W.H. Kane, E.W. Davie, Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders, Blood 71 (1988) 539–555.
[2] L.W. Hoyer, Hemophilia A, N. Engl. J. Med. 330 (1994) 38–47.
[3] United Kingdom Haemophilia Centre Doctors' Organization (UKHCDO), Guidelines on the selection and use of therapeutic products to treat haemophilia and other hereditary bleeding disorders, Haemophilia 9 (2003) 1–23.
[4] P.J. Fay, Factor VIII structure and function, Thromb. Haemost. 70 (1993) 63–67.
[5] E.W. Davie, K. Fujikawa, W. Kisiel, The coagulation cascade: initiation, maintenance, and regulation, Biochemistry 30 (1991) 10363–10370.
[6] G.A. Vehar, B. Koyt, D. Eaton, H. Rodriguez, D.P. O'Brien, F. Rotblat, H. Opperman, R. Keck, W.I. Wood, R.N. Harkins, E.G.D. Tuddenham, R.M. Lawn, D.J. Capon, Structure of human factor VIII, Nature 312 (1984) 337–342.
[7] D. Meyer, J.P. Girma, Von Willebrand factor: structure and function, J. Thromb. Haemost. 70 (1993) 99–104.
[8] M. Zaverio, M.D. Ruggeri, Structure of von Willebrand factor and its function in platelet adhesion and thrombus formation, Best Pract. Res. Clin. Haematol. 14 (2001) 257–279.
[9] A.J. Vlot, S.J. Koppelman, B.N. Bouma, J.J. Sixma, Factor VIII and von Willebrand Factor, Thromb. Haemost. 79 (1998) 456–465.
[10] J.A. Koedam, J.C. Meijers, J.J. Sixma, B.N. Bouma, Inactivation of human factor VIII By actived protein C. Cofactor activity of protein S and protective effect of von Willebrand factor, J. Clin. Invest. 82 (1988) 1236–1243.
[11] A.B. Federici, The factor VIII/von Willebrand factor complex: basic and clinical issues, Haematologica 88 (2003) EREP02.
[12] E.L. Saenko, D. Scandella, The acidic region of Factor VIII light chain and the C2 domain together form the high affinity binding site for von Willebrand Factor, J. Biol. Chem. 272 (1997) 18007–18014.
[13] D. Josic, H. Schwinn, M. Stadler, A. Strancar, Purification of factor VIII and von Willebrand factor from human plasma by anion-exchange chromatography, J. Chromatogr. B: Biomed Appl. 662 (1994) 181–190.
[14] K. Pock, A. Jungbauer, C. Kannicht, K. Loster, A. Buchacher, D. Josic, Electrophoretic analyses of clotting factor VIII concentrates, Anal Chim. Acta 372 (1998) 219–229.
[15] R. Adamson, Design and operation of a recombinant mammalian cell manufacturing process for rFVIII, Ann. Hematol. 68 (1994) S9.
[16] D. Scandella, Epitome specificity and inactivation mechanisms of factor VIII inhibitor antibodies, Vox Sang. 77 (1999) 17–20.
[17] K. Pock, A. Rizzi, D. Josic, Use of high-resolution techniques for the characterization of clotting factor VIII, J. Chromatogr. A 852 (1999) 175–188.
[18] N. Bihoreau, J.F. Veillan, C. Ramon, J.M. Scohyers, J.M. Schmitter, Characterization of a recombinant antihaemophilia-A factor (factor VIII-delta II) by matrix-assisted laser desorption/ionization mass spectrometry, Rapid Commun. Mass Spectrom. 9 (1995) 1584–1588.
[19] K.F. Medzihradsky, M.J. Besman, A. Burlingame, Reversed-phase capillary high performance liquid chromatography/high performance electrospray ionization mass spectrometry: an essential tool for the characterization of complex glycoprotein digests, Rapid Commun. Mass Spectrom. 12 (1998) 472–478.
[20] D.A. Wolters, M.P. Washburn, J.R. Yates III, An automated multidimensional protein identification technology for shotgun proteomics, Anal. Chem. 73 (2001) 5683–5690.
[21] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates III, Direct analysis of protein complexes using mass spectrometry, Nat. Biotechnol. 17 (1999) 676–682.
[22] H. Liu, R.G. Sadygov, J.R. Yates III, A model for random sampling and estimation of relative protein abundance in shotgun proteomics, Anal. Chem. 76 (2004) 4193–4201.
[23] L. Florens, M.P. Washburn, J.D. Raine, R.M. Anthony, M. Grainger, J.D. Haynes, J.K. Moch, N. Muster, J.B. Sacci, D.L. Tabb, A.A. Witney, D. Wolters, Y. Wu, M.J. Gardner, A.A. Holder, R.E. Sinden, J.R. Yates, D.J. Carucci, A proteomic view of the *Plasmodium falciparum* life cycle, Nature 419 (2002) 520–526.
[24] J.E. Addiego Jr., E. Gomperts, S.L. Liu, P. Bailey, S.G. Courter, M.L. Lee, G.G. Neslund, H.S. Kingdon, M.J. Griffith, Treatment of hemophilia A with a highly purified factor VIII concentrate prepared by anti-FVIIIc immunoaffinity chromatography, Thromb. Haemost. 67 (1992) 19–27.
[25] P.L. Mauri, A. Scarpa, A.C. Nascimbeni, L. Benazzi, E. Parmagnani, A. Mafficini, M. Della Peruta, C. Bassi, K. Miyazaki, C. Sorio, Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers, FASEB J. 19 (2005) 1125–1127.
[26] J.G. Clifton, F. Huang, S. Kovac, X. Yang, D.C. Hixson, D. Josic, Proteomic characterization of plasma-derived clotting factor VIII-von Willebrand factor concentrates, Electrophoresis 30 (2009) 3636–3643.
[27] P.L. Mauri, G. Dehò, A proteomic approach to the analysis of RNA degradosome composition in *Escherichia coli*, Methods Enzymol. 447 (2008) 99–117.
[28] J. Gao, G.J. Opiteck, M.S. Friedrichs, A.R. Dongre, S.A. Hefta, Changes in the protein expression of yeast as a function of carbon source, J. Proteome Res. 2 (2003) 643–649.
[29] S. Stoilova-McPhie, B.O. Villoutreix, K. Mertens, G. Kemball-Cook, A. Holzenburg, 3-Dimensional structure of membrane-bound coagulation factor VIII: modeling of the factor VIII heterodimer within a 3-dimensional density map derived by electron crystallography, Blood 99 (2002) 1215–1223.
[30] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, Quantitative mass spectrometry in proteomics: a critical review, Anal. Bioanal. Chem. 389 (2007) 1017–1031.
[31] S.K. Park, J.D. Venable, T. Xu, J.R. Yates III, A quantitative analysis software tool for mass spectrometry-based proteomics, Nat. Methods 5 (2008) 319–322.
[32] S.A. Gerber, J. Rush, O. Stemman, M.W. Kirschner, S.P. Gygi, Absolute quantification of proteins and phosphoproteins from call lysates by tandem MS, PNAS 100 (2003) 6940–6945.
[33] T.W. Barrowcliffe, Factor VIII and factor IX Sub-Committee. Recommendations for the assay of high-purity factor VIII concentrates, Thromb. Haemost. 70 (1993) 876–877.
[34] M. Wadhwa, M. Dilger, J. Tubbs, A. Mire-Sluis, T.W. Barrowcliffe, R. Thorpe, Identification of transforming growth factor-beta as a contaminant in factor VIII concentrates: a possible link with immunosuppressive effects in haemophiliacs, Blood 84 (1994) 2021–2030.
[35] T.C. Abshire, H.H. Brackmann, I. Scharrer, K. Hoots, C. Gazengel, J.S. Powell, E. Gorina, E. Kellermann, E. Vosburgh, Sucrose formulated recombinant human antihemophilic factor VIII is safe and efficacious for treatment of hemophilia A in home therapy—International Kogenate-FS Study Group, Thromb. Haemost. 83 (2000) 811–816.