CIEF AND MALDI-TOF-MS METHODS FOR ANALYZING FORMS OF THE GLYCOPROTEIN VEGF$_{165}$

Sara Ongay[1], Angel Puerta[2+], Jose Carlos Díez-Masa[1], Jonas Bergquist[2], Mercedes de Frutos[1]*

1. Institute of Organic Chemistry (C.S.I.C.). Madrid (Spain)

2. Analytical Chemistry, Department of Physical and Analytical Chemistry, Biomedical Centre, Uppsala University. Uppsala (Sweden)

+ Present address: Institute of Organic Chemistry (C.S.I.C.). Madrid (Spain)

*** Corresponding author: Mercedes de Frutos**

Institute of Organic Chemistry (C.S.I.C.).

Juan de la Cierva 3. 28006 Madrid (Spain)

Phone: +34 91 2587511

FAX: +34 915644853

E-mail: mfrutos@iqog.csis.es

Abbreviations:

IS, internal standard**; MW**, molecular weight; $\mathbf{t_m}$, migration time; **VEGF,** Vascular endothelial growth factor**; PA,** polyacrylamide.

KEYWORDS: Angiogenesis, Biomarker, Glycoprotein, Isoform, VEGF.

ABSTRACT

The vascular endothelial growth factor (VEGF) is involved in different sicknesses (cardiovascular diseases, cancer, and other). Out of the many components of the VEGF family, the A splice variant with 165 amino acids ($VEGF_{165}$) is the main component. In spite of the potential as biomarker that this protein has, information about its physico-chemical characteristics is scarce.

In this study CIEF and MALDI-TOF-MS methods for intact recombinant human $VEGF_{165}$ are developed and applied to analyze this glycoprotein expressed in glycosylating (*Sf 21* insect cells) and non-glycosylating (*E. coli*) systems. Different parameters influencing the CIEF separation were studied. The developed CIEF method allowed for the separation of up to seven peaks in the $VEGF_{165}$ expressed in insect cells and up to three in $VEGF_{165}$ expressed in *E. coli*. The use of the presented method permits the estimation of the apparent p*I* of the different forms of $VEGF_{165}$ expressed in insect cells to be in a range of 6.8 to 8.2. The three peaks with intermediate p*I* values are observed in the protein expressed in both systems, insect cells and *E. coli*. The MALDI-TOF-MS method enabled to a rapid partial characterization of $VEGF_{165}$ based on its MS fingerprint. MALDI-MS analysis of $VEGF_{165}$ expressed in insect cells shows the presence of, at least, 4 forms or groups of forms of $VEGF_{165}$ as a result of the different PTMs of the protein. According to MALDI-MS analysis, $VEGF_{165}$ expressed in *E. coli* was produced as a very homogeneous protein although the results suggest the existence of some PTMs in the protein. The patterns of $VEGF_{165}$ of both origins obtained by CIEF and MALDI-MS indicate the possibility of using these analytical methods to compare samples from people with different pathophysiological conditions. This work is thus a starting point to make possible the study of the role of the various forms of $VEGF_{165}$ as

biomarkers. Finally, to the best of our knowledge, this is the first time that intact $VEGF_{165}$ has been analyzed by CIEF and MALDI-TOF-MS.

1. INTRODUCTION

Glycoproteins play an important role in many pathophysiological processes. For this reason some of the compounds that are considered to be either actual or potential biomarkers are glycoproteins. Also, for the same reason, some glycoproteins are used as therapeutic drugs.

The vascular endothelial growth factor (VEGF), which actually is a family of glycoproteins (VEGF A, VEGF B, etc.), is known to have angiogenic, mitogenic, and vascular permeability enhancing effects [1-3]. Out of the many components of the family, the A splice variant with 165 amino acids, the so-called VEGF A 165, VEGF 165, or $VEGF_{165}$, is the main component.

Information about the physico-chemical characteristics of VEGF is scarce [1,4-6]. The human $VEGF_{165}$ is known to be a homodimeric, disulfide linked glycoprotein with an N-glycosylation site at Asn 75 [1, 7, 8].

Due to the fact that the different isoforms of VEGF have differing functions, changes in the isoform pattern without change in the total VEGF concentration could have complex effects [9] and have an impact on processes such as carcinogenesis [10].

Different studies concerning the role of VEGF as a biomarker have been carried out [11-13]. Although $VEGF_{165}$ is the predominant and most angiogenic form of the glycoprotein, very few studies about its role as a biomarker have been carried out [14].

All the heterogeneities of a protein, due to any PTM event or to the existence of genetic variants, are generally termed "forms" of a protein, while forms arising from differences in the glycosidic part of the protein are specifically called "glycoforms". As it is the case for other glycoproteins, more than one form of $VEGF_{165}$ can exist. Similarly to what occurs for other glycoproteins [15, 16] the ratio of some of these forms might have a better value as a biomarker than the whole level of the glycoprotein. Besides, as it is the case for other therapeutical glycoproteins [17], the biological activity of the protein drug may differ from form to form. For these reasons, distinguishing between different forms of $VEGF_{165}$ could be of interest.

In a previous study carried out in our laboratories, two different CE methods were developed to analyze $VEGF_{165}$ expressed in glycosylating (*Sf21* insect cells) and non-glycosylating (*E. coli*) systems [18]. The existence of different forms of the glycoprotein was observed.

Differences in charge and/or in size can arise from changes in glycosylation or from other PTMs such as phosphorylations, the removal of one or more amino acids from the amino terminus, or amino acid oxidation. Changes in the amino acid sequence of $VEGF_{165}$ corresponding to cleavage of some amino acids in the N-terminus or oxidation of methionine to methionine sulfoxide have been previously described [19].

In isoelectric focusing molecules are separated according to their isoelectric point, being then a technique appropriate to distinguish molecules bearing different charge. The capillary version of IEF, CIEF, presents advantages compared to IEF in the slab gel format. It is faster, more reproducible, quantitative results are obtained in the capillary

format, and the amount of ampholytes required for the analysis is smaller than in the slab gel format [20].

MALDI-MS is one of the main MS technologies for the analysis of peptides and proteins. As MALDI is a soft ionization technique, mainly singly charged molecules are produced and it makes possible to measure the molecular weight (MW) of the ionized analyte. As a consequence, this technique enables to distinguish molecules based on mass differences and to provide the mass fingerprint of the sample. The above mentioned production of mostly singly charged ions is a remarkable advantage in the analysis of complex samples such as a mixture of different forms of a glycoprotein, the objective of this work. Other advantages of MALDI-MS are high sensitivity, short analysis time, spectral simplicity, relatively high tolerance towards non-volatile buffers, and detection of high mass molecules when TOF analyzer is used [21, 22].

In this study CIEF and MALDI-MS methods to analyze recombinant human $VEGF_{165}$ were developed and applied to analyze this glycoprotein expressed in glycosylating (*Sf 21* insect cells) and non-glycosylating (*E. coli*) systems. The methods should allow us to gain knowledge about the origin of the differences observed by our previous CE studies [18] for these samples. Furthermore, these methods, by permitting comparison between $VEGF_{165}$ from different samples, could be the basis for performing studies concerning the role of the different forms of $VEGF_{165}$ as biomarkers of different diseases. Additionally, these methods should allow carrying out quality control of the glycoprotein if used as a therapeutic drug.

2. MATERIALS AND METHODS

2.1. Chemicals, solutions, and samples

Recombinant Human Vascular Endothelial Growth Factor 165 ($VEGF_{165}$), (293-VE/CF) expressed in *Sf* 21 insect cells was obtained from R&D Systems Europe Ltd (Abingdon, UK). Non-glycosylated recombinant DNA-derived $VEGF_{165}$ expressed in *E. coli* cells (code 01/424) was purchased from NIBSC (Hertfordshire, UK).

Low molecular weight (MW) components of the $VEGF_{165}$ solutions were eliminated employing centrifugal filter devices with Mr 10000 cut-off membrane (Microcon® 10, Millipore). Previously to the use of the centrifugal filter devices, they were passivated with a 5 % Brij® 35 solution (w/v) to avoid the non-specific adsorption of the proteins on the plastic walls of the filtration device [23]. Recovery of $VEGF_{165}$ was not measured and thus, the concentrations mentioned along the text are approximated values.

Further information about chemicals and solutions is available in the Supporting Information.

2.2. Equipments

CIEF separations were carried out in a PACE 5000 CE instrument (Beckman, Instruments, Fullerton, CA, USA). Data were collected and analyzed using the System Gold software (version 8.1) from Beckman.

MALDI-TOF mass spectra were recorded on a Bruker Ultraflex II MALDI-TOF/TOF instrument (Bruker Daltonics, Bremen, Germany). The software packages Ultraflex

version 3.0 and Flex Analysis version 3.0 were used to record and analyze mass spectra, respectively.

2.3. CIEF procedures

Two different types of capillaries were used: Polyvinylalcohol (PVA)-coated capillaries (N-CHO coated capillary from Beckman Coulter) and polyacrylamide (PA)-coated capillaries (Neutral capillary from Beckman Coulter). In both cases, the capillary id was 50 µm and total length was 27 cm with an effective length of 20 cm. Temperature was set at 20 ºC and detection was performed at 280 nm. Conditioning of each new PVA-coated or PA-coated capillary was made by rinsing it (N_2 pressure 13.8 bar) with 10 mM H_3PO_4 for two minutes and water for 10 minutes. Capillaries were rinsed between injections (using 13.8 bar N_2 gas) with 10 mM H_3PO_4 for 1 minute, water for 2 minutes, and CIEF gel for 3 minutes. To be stored capillaries were rinsed for 10 minutes with water, both ends were immersed in water, and they were kept at 4 ºC.

The effect of different concentrations of H_3PO_4 and NaOH as anolyte and catolyte was studied during the CIEF method development. The mixture of ampholytes (optimized as shown below), urea (if any), and NaCl in different concentrations (when added) in CIEF gel is called separation mixture along this study. Different proteins, such as ribonuclease B (rib, pI 9.3), horse myoglobin (myo, pI 7.3 and 6.9) and β lactoglobulin A (lg, pI 5.2), were studied as possible internal standards (IS).

The way of performing the injection of the sample was optimized as shown below. Focusing voltage and focusing time were also optimized. Hydrodynamic mobilization

was performed after the focusing step by applying N_2 pressure (34 mbar) at the anodic end while voltage was maintained.

2.4. $VEGF_{165}$ peak assignment. Computer program and migration parameters

Migration behavior of CE peaks is usually characterized by the migration time (t_m). We should clarify that even though talking about t_m is not strictly correct in CIEF we will employ this term to facilitate comprehension. The t_m values will correspond to the time elapsed after injection to the detection of the peak. This parameter depends in CIEF on many experimental conditions and it is difficult to reproduce. In order to compare protein peaks of different samples, more robust migration parameters are needed. The usefulness of a computer program developed in our laboratory for another glycoprotein [24] was tested in this work to find a suitable migration parameter to correctly assign peaks of $VEGF_{165}$ separated by CIEF. The migration parameters considered in this work were: P1 (migration time, t_m), P2 (migration time relative to peak 1 of myoglobin, t_m/t_{myo1}), P3 (migration time relative to peak 2 of myoglobin, t_m/t_{myo2}), P4 (migration time relative to ribonuclease B, t_m/t_{rib}), P5 (migration time relative to β lactoglobulin, t_m/t_{lg}), and P6 (migration time relative to anolyte front, t_m/t_{front}). As we are employing pressure mobilization the anolyte is introduced and pushed in the capillary reaching the detector and giving rise to an absorbance front at the end of the electropherogram. Further details about the way of performing this study are available in the Supporting Information.

2.5. MALDI-TOF-MS measurements

MALDI-TOF-MS spectra of human $VEGF_{165}$ expressed either in insect cells or in *E. coli* were collected using linear acquisition operation mode with 25 kV acceleration

voltage. Samples were applied on the MALDI target using the sinapinic acid double layer method [25, 26]. Shortly, a thin layer of ethanol saturated with sinapinic acid was applied on the target. Afterwards, 1 µL of the $VEGF_{165}$ sample was mixed with 1 µL of 30 % ACN + 0.1 % trifluoracetic acid (TFA) solution saturated with sinapinic acid and 1 µL of this mixture was applied on the target. $VEGF_{165}$ samples were prepared either by diluting the commercial sample in 0.1 % TFA or by filtering it with ultracentrifuge filter devices to eliminate the low molecular weight components and reconstituting it in water.

3. RESULTS AND DISCUSSION

3.1 CIEF

3.1.1 Sample preparation and injection

The most widespread way of working in CIEF consists of injecting the sample by blending it with a mixture of ampholytes dissolved in a specific polymer for CIEF that can contain in some instances substances that enhance solubility or denaturing agents. This mixture containing the sample is introduced in the capillary by employing high pressure; then, focusing and mobilization are performed. However, this sample injection procedure is not very suitable for the development of the CIEF method to analyze $VEGF_{165}$ peaks due to the high sample consumption for this expensive protein. For this reason, the possibility of injecting the protein without adding it to the separation mixture has been studied [27, 28].

To select the most appropriate form of preparing and injecting the sample without using the expensive $VEGF_{165}$, some model proteins were employed instead of $VEGF_{165}$. In this preliminary study the separation mixture consisted on 2% ampholytes of pH range

3-10 in CIEF gel, 7 M urea, and 10 mM sodium chloride. As sample, a mixture of proteins consisting on ribonuclease B (p*I* 9.3), myoglobin (p*I* 7.3 and 6.9), and β lactoglobulin (p*I* 5.2) was employed. These proteins were chosen as their p*I* values should cover the expected p*I* range of the $VEGF_{165}$ peaks, although the p*I* value of 8.5 given in the literature [4] is for VEGF purified from bovine pituitary and not for the human protein. A 20 kV focusing voltage during 3 min was employed as the focusing step. Five different sample injection procedures were compared. Details about the way of performing them and the results obtained are available in Supporting Information.

Based on these results and the high consumption of protein when normal injection was used, high pressure mode, consisting on introducing consecutively a large plug of the separation mixture and a small plug of the protein at high pressure (13.8 bar), was chosen as the most appropriate one for the development of a CIEF method for $VEGF_{165}$ peak separation.

3.1.2 Selection of catholyte and anolyte

There are two effects, anodic and cathodic drift, that can effect CIEF separations leading to the loss of the analyte or analytes of interest. These effects can be controlled by the nature of catholyte and anolyte in the separation. As it has been described by Mosher et al. [29] for slab gel IEF, anodic and cathodic drifts are pH instabilities that cause a progressive loss of the acid and basic end of the pH gradient. If the protein focuses close to the ends of the gradient some of the peaks can be lost due to these effects. To see if anodic or cathodic drift effected the protein analysis, different combinations of catholyte and anolyte were studied analyzing recombinant human $VEGF_{165}$ expressed in insect cells employing the same separation mixture and focusing

conditions as in the previous step. H_3PO_4 dissolved in CIEF gel at two different concentrations (98 mM and 20 mM) were employed as anolyte and NaOH dissolved in water at two concentrations (100 mM and 20 mM) were employed as catholyte and the results were compared. No differences in VEGF peak separation could be observed for any of the four combinations of anolyte and catholyte tested. This would indicate that under these conditions neither cathodic nor anodic drift are effecting the protein separation, possibly because the protein is not focusing close to the very ends of the pH gradient.

3.1.3 Effect of the composition of the separation mixture

The effect of salts in the separation mixture in cIEF is constantly under discussion [30-34]. The effect of different concentrations of NaCl (5-60 mM) in the separation mixture was studied employing recombinant human $VEGF_{165}$ expressed in insect cells using the same separation mixture and focusing conditions as in the previous steps.

No effect in $VEGF_{165}$ peak separation was observed for the salt concentration interval studied. For this reason, salt addition was not employed in the subsequent CIEF method development.

It has been reported that for the separation of proteins with close p*I* values, the mixture of several ampholytes from different manufacturers reinforces the efficiency of the separation in the p*I* range of interest, giving rise to better resolution than when employing a single ampholyte [20, 35]. For this reason, different mixtures of ampholytes (with pH ranges 3-10, 5-8, 8-10 and 9-11) in different total percentage

(from 0.25 to 10%) and different ratios were studied in this work. The detailed way of performing the study and the results are available in the Supporting Information.

As a result of this study, 1% of ampholytes consisting of a mixture of pH ranges 5-8 and 8-10 in a 1:1 ratio was selected.

Urea is commonly used in CIEF for protein analysis as it can suppress protein aggregation, enhancing their solubility [36]. Urea concentrations in the range 6-8 M in separation mixture are recommended to avoid protein precipitation [37] and it has been described that the addition of urea to the separation mixture enhanced separation of peaks of glycoproteins as α-1-acid glycoprotein and erythropoietin by CIEF [20, 33, 38]. Thus, the influence of the presence of 7 M urea in the separation mixture was studied under the conditions optimized in the prior steps. No differences in protein peak resolution were observed although spikes were more abundant in absence of urea (data not shown). Taking into account that this method is being developed for future application to more complex samples, where protein precipitation could be anticipated due to higher protein concentration in the sample, urea was employed in the separation mixture.

3.1.4 Effect of voltage and focusing time

A succesful focusing step is of major importance in CIEF. An incomplete focusing at the time of detection can be responsible for loss in resolution [39]. Discrepancies exist regarding the length of focusing time [34, 37, 40]. Applied voltage and focusing time were optimized for the separation of human $VEGF_{165}$ expressed in insect cells.

Voltage was varied in the range 5 to 25 kV with a constant focusing time of 3 min. Migration time decreased when increasing the voltage. This fact could be attributed to incomplete focusing at low voltages and to remaining electroosmotic flow in both types of coated capillaries used, an effect that increased with capillaries aging. Improvement in resolution was observed when voltage was increased. For this reason 25 kV was chosen as the optimal focusing voltage for subsequent analysis. Different focusing times (0.2 – 25 min) were studied applying 25 kV focusing voltage (data not shown). Under these conditions (1% ampholytes of pH ranges 5-8 and 8-10 (1:1) dissolved in CIEF gel, 7 M urea) initial current value was 3 μA. Focusing time shorter than 3 min was not enough to obtain good resolution of the $VEGF_{165}$ peaks. An increase of focusing time in the interval from 3 to 10 min led to improvement of the resolution of the $VEGF_{165}$ peaks, and at the time of 10 min the current had just decreased below 20 % of its initial value. Thus, a focusing time of 10 min provided the best resolution values without any improvement when focusing time was further increased. It was also observed that the salts remaining in the sample (from the ultrafiltration cleaning step) effected the focusing of the protein; for low salt concentration a focusing time of 10 min was not enough to obtain the best resolution. For this reason, a 15 min focusing time was selected as optimum one to be sure that even with less salty samples the focusing step was optimum. Under these conditions, the ribonuclease was focused at the portion of capillary past the detection window, as indicated by showing a migration time lower than the focusing time. None of the $VEGF_{165}$ peaks was focused at this part of the capillary column.

3.1.5 Evaluation of capillary coating

As indicated in materials and methods (see section 2) two different types of capillaries, polyvinylalcohol (PVA)-coated and polyacrylamide (PA)-coated, were employed during the CIEF development. The influence of the type of capillary coating on the resolution of $VEGF_{165}$ peaks was tested. No differences in $VEGF_{165}$ peak resolution was observed as a function of the type of capillary coating, though PVA-coated capillaries led to longer migration times than the ones observed for PA-coated capillaries. Both types of capillaries were compared in different studies in order to follow their behavior with long-term use. Best behavior according to long-term use was observed for PA-coated capillaries.

In summary, the optimized CIEF method parameters to analyze $VEGF_{165}$ peaks were: PA-coated capillary (50 µm id, 20 cm effective length, 27 cm total length); consecutive high pressure injection mode of separation mixture (1% ampholytes of pH range 5-8 and 8-10 (1:1) dissolved in CIEF gel containing 7 M urea) for 2.4 min and protein sample for 0.2 min; 25 kV focusing voltage and 15 min focusing time followed by pressure mobilization (34 mbar N_2) while keeping the voltage; 98 mM phosphoric acid in CIEF gel as anolyte and 20 mM sodium hydroxide as catholyte; temperature at 20ºC and detection at 280 nm.

3.1.6 Assignment accuracy of $VEGF_{165}$ peaks

As indicated in section 2.4, six migration parameters were compared to characterize the migration of $VEGF_{165}$ peaks. A four-day study was performed to calculate the accuracy of assignment provided by each parameter for human $VEGF_{165}$ expressed in insect cells. Different $VEGF_{165}$ concentrations (1.6, 3.3 and 5.0 mg/mL) were consecutively injected with the other proteins used as internal standards and analyzed by the developed CIEF

method. Duplicate injections of each VEGF concentration were analyzed in an increasing (first day), decreasing (second day) and random concentration order (third day). On the fourth day, 6 injections of 1.6 mg/mL $VEGF_{165}$ were performed. This study was carried out co-injecting several proteins in order to know if the presence of different concentrations of VEGF could affect the reliability of the migration parameters selected for $VEGF_{165}$ peaks assignment. The rationale for performing this study was that different protein concentrations are expected in natural samples, and proteins can act as carrier ampholytes, altering the pH gradient [41]. The detailed results of this study are shown in Table 1 of Supporting Information.

It was observed that parameters P2 and P3 (migration time relative to peak 1 and to peak 2 of myoglobin, respectively) provided the best accuracy, close to 100%, in $VEGF_{165}$ peak assignment. No effect of $VEGF_{165}$ concentration in peak assignment was observed.

3.1.7 p*I* estimation of recombinant human$VEGF_{165}$ expressed in insect cells

In this work, the isoelectric points of the seven major peaks of recombinant human $VEGF_{165}$ expressed in insect cells, separated by the CIEF method proposed, were estimated employing standard proteins as p*I* markers. The calculated p*I* values were in the range 6.8 – 8.2. A more detailed description of the way of calculating the p*I* and Table 2 containing the individual values for each peak can be seen in Supporting Information. The calculated p*I* values should be considered as apparent values, because some factors could influence the p*I* value determined by this CIEF method [41]. For instance, the presence of urea in the separation mixture can most probably modify the values of p*I* [42]. The values presented for $VEGF_{165}$ (6.8-8.2) are lower than the value

of 8.5 given in the literature [4], reasonable taken into consideration that this value is for VEGF purified from bovine pituitary and not for the human one. In addition, the p*I* reported in literature should correspond to the average value of the different VEGF components calculated under other experimental conditions.

3.1.8 Comparison of samples of recombinant human $VEGF_{165}$ expressed in insect cells and *E. coli*

The CIEF method developed and the migration parameters selected were employed for the comparison of human $VEGF_{165}$ expressed in insect cells, which is glycosylated [43], and in *E. coli*, which is not glycosylated (Figure 1). $VEGF_{165}$ expressed in insect cells presented seven major peaks while $VEGF_{165}$ expressed in *E. coli* presented only three. The main peak for both proteins presented the same value of migration parameter P2 (one out of the two parameters providing close to 100% accuracy of assignment), while the most basic and acidic peaks (1, 2, 6, and 7) of $VEGF_{165}$ observed for insect cells were either not present or present at very low level in the $VEGF_{165}$ expressed in *E. coli*. Peaks 3, 4, and 5 for the protein expressed in insects have the same values of migration parameter P2 as have the peaks a, b, and c expressed in *E. coli*. However this correspondence in P2 values, and thus in p*I* values (see Table 2 in Supporting Information), does not imply that the composition of the protein is the same; there could be different forms of the protein with very similar p*I* average values under each of the peaks. Different appearance was observed for p*I* markers in both samples. This effect was repeatable and would most likely be due to remaining excipients from the samples that interferes with the focusing step. The observed differences in the focusing step of the two samples of $VEGF_{165}$ (expressed in insect and in *E. coli*.) makes the comparison intricate.

3.2. MALDI-TOF-MS measurements

In order to characterize the average size of the components of $VEGF_{165}$ samples, MALDI-TOF-MS measurements were performed. $VEGF_{165}$ expressed in insect cells was measured after either dilution with 0.1 % TFA to a concentration range of 0.01-0.5 mg/mL or after removing low molecular weight components from the samples by centrifugal filter devices, resulting in a theoretical concentration of the glycoprotein in the range of 0.2-0.6 mg/mL. Best results in the measurements of $VEGF_{165}$ expressed in insect cells were obtained by direct dilution of the sample as compared to the samples where low molecular components were removed by cut-off filter devices (data not shown). Most likely, traces of Brij® 35 employed in the pretreatment of the filter devices could be present in the sample applied on the MALDI target disturbing the formation of the crystals and promoting more noisy signals. It has been described in the bibliography that depending on the matrix, the protein, and the sample application method on the MALDI probe, the presence of contaminating salts, buffers or surfactants (ionic and non-ionic) can produce a decrease in the quality of the spectra [44-48]. Figure 2 shows the MALDI-TOF-MS spectrum of $VEGF_{165}$ expressed in insect cells including the main values of the mass peaks (37995.5, 38909.3, 39019.0, 39784.4). At least four main mass peaks, each of them with several components, which could correspond to either four main forms or four main groups of forms of $VEGF_{165}$ expressed in insect cells were observed. The mass differences between the peaks of the MS spectra cannot be explained by artifacts due to adduct formation with TFA molecules. Explanation of the different peaks regarding to PTMs (specially regarding to differences in the glycosylation of the forms) was not achieved due to the noise present in the measurements which resulted in less accurate mass determination.

In order to compare the mass peak pattern of recombinant human $VEGF_{165}$ expressed in a glycosylating system (insect cells) and in a non-glycosylating system (*E. coli*), the World Health Organization reference reagent for $VEGF_{165}$ which is expressed in *E. coli* was measured. Commercial preparation of $VEGF_{165}$ expressed in *E. coli* contained trehalose in a concentration 3000 times higher than the concentration of $VEGF_{165.}$ The excess of trehalose in the sample could disturb the formation of the matrix crystals and as a consequence only samples where the low molecular weight components were removed by cut-off filter devices were analyzed. Figure 2 shows the MALDI-TOF-MS spectrum of $VEGF_{165}$ expressed in *E. coli*. The molecular weight of $VEGF_{165}$ expressed in *E. coli* according to these results is 38179.8 Da. The observation of a single peak indicates that $VEGF_{165}$ expressed in *E. coli* is rather homogenous, or that there are several isoforms with almost the same mass, as a result of different PTMs of similar mass shifts (reduction of a disulfide bridge *vs* deamidation of asparagine or glutamine; hydroxylation *vs* methionine oxidation *vs* methylation).

According to the structural information in the Expasy database, the theoretical molecular weight of the $VEGF_{165}$ homodimer without any PTM is 38316.2 Da [8], though some authors [19] suggest that the real MW of the protein is 38300.1 Da. Even with less accurate mass determination in the MALDI-TOF-MS measurements, there is a significant disagreement between the theoretical MW of $VEGF_{165}$ (either 38316.2 or 38300.1 Da) and the measured MW of $VEGF_{165}$ expressed in *E. coli*. This could indicate that there are some PTMs or degradation processes involving cleavage at the amino acid sequence. This mass difference could be due to loss of one or two amino acids from the primary structure of the protein, either arginine from the C-terminus from

one monomer or alanine from the N-terminus [19] of both monomers. In the same way, in the MALDI-MS spectrum of $VEGF_{165}$ expressed in insect cells shown in Fig. 2, it is possible to observe a first peak with mass value near 38000 Da, a value lower than the one expected for the amino acid sequence of $VEGF_{165}$, suggesting that some of the components or forms of $VEGF_{165}$ expressed in insect cells could be produced by some truncation in the amino acid sequence and not only due to glycosylation processes that result in forms with higher mass values. The mass difference between the 1st peak in the spectrum of $VEGF_{165}$ expressed in insect cells (Fig. 2, black line) and the theoretical MW for the amino acid sequence of $VEGF_{165}$ could match with the loss of alanine-proline-methionine from the N-terminus of one monomer [19] or the loss of the arginine residue from the C-terminus of both monomers, though it could be a more complex process such as glycosylation and truncation of several amino acids in combination.

Although it is not possible to fully explain the PTMs in $VEGF_{165}$ expressed in insect cells based on these results, it is remarkable that the mass of these glycoforms are in the range 38000-40100 Da. This mass range disagrees with the reported MW of glycosylated $VEGF_{165}$ of 42-46 kDa [4,6,19,49]. However, either these MW values were measured by SDS-PAGE analyses which are not very accurate in mass determination, or the analyzed protein was expressed in an organism different to the one employed to produce $VEGF_{165}$ analyzed in this work [19, 49] which usually means different glycosylation of the protein.

Comparison of results of $VEGF_{165}$ analysis by CIEF and MALDI-MS methods and by previously published CE methods developed in our laboratories [18] is not straight forward. MALDI-MS analysis of $VEGF_{165}$ expressed in *E. coli* showed the protein as a

single mass peak, but when the same protein was analyzed by CE, it migrated as either 2 (CE analysis with bare silica capillaries) or 3 (CIEF and CZE analysis with coated silica capillaries) peaks. PTMs such as deamidation of asparagine or glutamine to aspartic or glutamic acid, which hardly modify the mass of the molecule, could produce a shift in the p*I* of the protein and as a consequence could be responsible of the discrepancies observed between MALDI-MS and CIEF results. Loss of amino acids from the N-terminus of $VEGF_{165}$ as earlier reported [19] did not probably contribute to the observance of different peaks in the CIEF analysis of $VEGF_{165}$ expressed in *E. coli* as those amino acids cannot hold any charge at their R-residues. The mentioned loss of amino acids would imply the presence of a different amino acid at the N-terminus whose pK of the $-NH_3^+$ group would be slightly different, meaning that there could be a small shift in the p*I* of the whole protein. However, it is not clear if this shift could have a significant effect on the total p*I* value. The MALDI-MS analysis of $VEGF_{165}$ expressed in insect cells showed the presence of 4 different mass peaks, though when this protein was analyzed by CE, either 5 (CE with bare silica capillaries) or 7 peaks (CIEF and CZE with coated silica capillaries) were observed. Peaks observed at MALDI-MS analysis could be promoted by PTMs involving loss of amino acids from the N-terminus and/or changes in the glycosylation. Peaks observed in the CIEF analysis could be due to the deamidation processes and changes in the glycosylation of the protein involving variation in the number of charged carbohydrates, such as sialic acid, which will modify the p*I* of the glycosylated protein. Change in the number of neutral carbohydrates could be shown in the MALDI-MS spectra though they would not be observed in the CIEF analysis.

4. CONCLUDING REMARKS

CIEF and MALDI-MS methods for the analysis of recombinant human $VEGF_{165}$ expressed in different cells, without any prior hydrolysis step, have been developed. Different parameters influencing the CIEF separation were studied in order to achieve a method that allowed the separation of up to seven peaks in the VEGF expressed in insect cells and up to three in the VEGF expressed in *E. coli*. Employing this method an estimation of the p*I* values for the peaks of $VEGF_{165}$ expressed in insect cells was performed. Results obtained indicate that the apparent p*I* of the different forms of $VEGF_{165}$ under these experimental conditions is in the ranges of 6.8 to 8.2 and 7.5 to 7.9 for the protein expressed in insect cells and expressed in *E. coli*, respectively. The described MALDI-TOF-MS method enables to partially characterize $VEGF_{165}$ expressed in different organisms providing their MS fingerprint in a very short time in comparison with the methods involving hydrolysis of the protein and a separation technique additional to the MS analysis. MALDI-MS analysis of $VEGF_{165}$ expressed in insect cells shows the presence of, at least, 4 forms or groups of forms of $VEGF_{165}$ due to the PTMs of the protein. According to MALDI-MS analysis, $VEGF_{165}$ expressed in *E. coli* was produced as a very homogenous protein, though mismatching between the measurements and the theoretical MW of the protein suggests the existence of some PTM in the protein. To the best of our knowledge, this is the first time that intact $VEGF_{165}$ has been analyzed by MALDI-TOF-MS and CIEF showing the complexity of the protein due to its PTMs.

Acknowledgements

Sara Ongay acknowledges C.S.I.C for a predoctoral grant. Angel Puerta acknowledges Spanish Ministry of Education and Science for a postdoctoral grant to perform research in Sweden. Jörg Hanrieder is acknowledged for help with the MALDI-TOF-MS. Financial support from the Spanish Ministry of Education and Science (Project CTQ2006-05214), Fundación Ramón Areces, Comunidad de Madrid (Project S-GEN/0247/2006) and the Swedish Research Council (projects 621-2005-5279, 629-2002-6821 J.B) is acknowledged. Beckman Coulter is acknowledged for providing some of the CE capillaries.

5. REFERENCES

[1] Leung, D.W., Cachianes, G., Kuang, W.-J., Goeddel, D.V. et al., *Science*, 1989, *246*, 1306-1309.

[2] Ferrara, N., Gerber H.-P., LeCouter, J., *Nat. Med.* 2003, *9*, 669-676.

[3] Ferrara, N., *Endocr. Rev.* 2004, *25*, 581-611.

[4] Ferrara, N., Houck, K.A., Jakeman, L.B., Winer, J., Leung, D.W., *J. Cell. Biochem.* 1991, *47*, 211-218.

[5] Robinson, C.J., Das, R.G., Stammers, R., Rafferty, B., *Growth Factors* 2006, *24*, 285-290.

[6] Lee, G.Y., Jung, W.W., Kang, C.S., Bang, I.S., *Protein Expr. Purif.* 2006, *46*, 503-509.

[7] Yeo, T.K., Senger, D.R., Dvorak, H.F., Freter, L., Yeo, K.T., *Biochem. Biophys. Res. Commun.* 1991, *179*, 1568-1575.

[8] www.expasy.org/uniprot/P15692, August 26, 2008.

[9] Mathieson, P.W., *J. Am. Soc. Nephrol.* 2006, *17*, 602-603.

[10] Catena, R., Muniz-Medina, V., Moralejo, B., Javierre, B., et al., *Int. J. Cancer* 2007, *120*, 2096-2109.

[11] Adams, J., Carder, P.J., Downey, S., Forbes, M.A., et al., *Cancer Res.* 2000, *60*, 2898-2905.

[12] Maeda, K., Chung, Y.S., Ogawa, Y., Takaksuka, S., et al., *Cancer* 1996, *77*, 858-863.

[13] Suzuki, H., Murakami M., Shoji, M., Iso, Y., et al., *Coronary Art. Dis.* 2003, *14*, 301-307.

[14] Elting, J.J., Carney, W.P., Hamer, P.J., Patent WO200747955-A2.

[15] Girard, M., Lacunza, I., Diez-Masa, J.C., de Frutos, M., in Landers, J.P. (Ed.), *Capillary and Microchip Electrophoresis and Associated Microtechniques*, CRC Press, Boca Raton, FL 2008, pp 631-705.

[16] Lacunza, I., Kremmer, T., Diez-Masa, J.C., Sanz, J., de Frutos, M., *Electrophoresis* 2007, *28*, 4447- 4451.

[17] Egrie, J.C., Browne, J.K., *Nephrol. Dial. Transplant.* 2001, *16 (suppl 3)*, 3-13.

[18] Ongay, S., Puerta, A., Díez-Masa, J.C., J. Bergquist, J., de Frutos, M., *Electrophoresis*, "in press".

[19] Keck, R.G., Berleau, L., Harris, R., Keyt, B.A., *Arch. Biochem. Biophys.* 1997, *344*, 103-113.

[20] Cifuentes, A., Moreno-Arribas, M.V., de Frutos, M., Diez-Masa, J.C., *J. Chromatogr. A* 1999, *830*, 453-463.

[21] Nielen, M.W.F., *Mass Spectrom. Rev*. 1999, *18*, 309-344.

[22] Schlosser, G., Pocsfalvi, G., Malorni, A., Puerta, A., de Frutos, M., Vekey, K., *Rapid Commun. Mass Spectrom.* 2003, *17*, 2741-2747.

[23] Puerta, A., Diez-Masa, J.C., de Frutos, M., *J. Immunol. Methods* 2004, *89*, 225-237.

[24] Lacunza, I., Lara-Quintanar, P., Moya, G., Sanz, J., et al. *Electrophoresis* 2004, *25*, 1569-1579.

[25] Xiang, F., Beavis, C., *Rapid Commun. Mass Spectrom.* 1994, *8*, 199-204.

[26] Dai, Y., Whittal, R.M., Li, L., *Anal. Chem.* 1996, *68*, 2494-2500.

[27] Hempe, J.M., Craver, R.D., *Clin. Chem.* 1994, *40*, 2288-2295.

[28] Chen, S-M., Wiktorowicz, J.E., *Anal. Biochem.* 1992, *206*, 84-90.

[29] Mosher, R.A., Thornmann, W., *Electrophoresis* 1990, *11*, 717-723.

[30] Mazzeo, J.R., Krull, I.S., *J. Chromatogr.* 1992, *606*, 291-296.

[31] Shimura, K., *Electrophoresis* 2002, *23*, 3847-3857.

[32] Clarke, N.J., Tomlinson, A.J., *Anal. Chem.* 1997, *69*, 2786-2792.

[33] Wu, J., Li, S-C., Watson, A., *J. Chromatogr. A* 1998, *817*, 163-171.

[34] Lopez-Soto-Yarritu, P., Díez-Masa, J.C., Cifuentes, A., de Frutos, M., *J. Chromatogr. A*, 2002, *968*, 221-228.

[35] Wehr, T., Zhu, M., Rodríguez-Díaz, R., *Methods Enzymol.* 1996, *270*, 358-374.

[36] Rodríguez-Díaz, R., Wehr, T., Zhu, M., *Electrophoresis* 1999, *830*, 453-463.

[37] Righetti, P.G., Gelfi, C., Conti, M., *J. Chromatogr. B* 1997, *699*, 91-104.

[38] Lacunza, I., Díez-Masa, J.C., de Frutos, M., *Electrophoresis* 2007, *28*, 1204-1213.

[39] Molteni, S., Frischknecht, H., Thornman, W., *Electrophoresis* 1994, *15*, 22-30.

[40] Wu, J., Huang, T., Pawliszyn, J., in Landers, J.P. (Ed.), *Capillary and Microchip Electrophoresis and Associated Microtechniques*, CRC Press, Boca Raton, FL 2008, pp 563-580.

[41] Righetti, P.G., *J. Chromatogr. A* 2004, *1037*, 491-499.

[42] Ui, N., *Biochim. Biophys. Acta* 1971, *229*, 567-581.

[43] Lee, G.Y, Jung, W.W., Kang, C.S., Bang, I.S, *Protein Expr. Purif.* 2006, *46*, 503-509.

[44] Zaluzec, E.J., Gage, D.A., Watson, J.T., *Protein Expr. Purif.* 1995, *6*, 109-123.

[45] Schnaible, V., Michels, J., Zeth, K., Freigang, J., et al., *Int. J. Mass Spectrom. Ion Process*. 1997, *169/170*, 165-177.

[46] Karas, M., Bahr, U., Ingendoh, A., Nordhoff, E., et al., *Anal. Chim. Acta* 1990, *241*, 175-185.

[47] Zhang, N., Li, L., *Rapid. Commun. Mass Spectrom*. 2004, *18*, 889-896.

[48] Amado, F.M.L., Santana-Marques, M.G., Ferrer-Correia, A.J., Tomer, K.B., *Anal. Chem*. 1997, *69*, 1102-1106.

[49] Ferrara. N., Leung, D.W., Cachianes, G., Winer, J., Henzel, W.J., *Methods Enzymol*. 1991, *198*, 391-404.

FIGURE LEGENDS

Fig. 1.- CIEF electropherograms of human $VEGF_{165}$ expressed in A) insect cells, and B) *E. coli* cells. Polyacrylamide-coated capillary: 50 µm id, 20 cm effective length, 27 cm total length. Anolyte: 98 mM phosphoric acid in CIEF gel. Catholyte: 20 mM sodium hydroxide. Separation mixture: 1% (v/v) ampholytes of pH ranges 5-8 and 8-10 (1:1), 7 M urea, CIEF gel. Co-injection of $VEGF_{165}$ and internal standards. Focusing step: 15 min at 25 kV. Mobilization step: 25 kV and 34 mbar N_2 pressure. Detection: 280 nm. Temperature: 20ºC. Peak identification: (rib) ribonuclease B, (myo_1 and myo_2) horse myoglobin peaks, (lg) β lactoglobulin A, (1-7) peaks corresponding to human $VEGF_{165}$ expressed in insect cells, (a-c) peaks corresponding to $VEGF_{165}$ expressed in *E. coli* cells.

Fig. 2.- MALDI-TOF-MS spectra of 0.5 mg/mL $VEGF_{165}$ expressed in insect cells in 0.1 % trifluoroacetic acid solution (black line) and 0.08 mg/mL $VEGF_{165}$ expressed in *E. coli* diluted in water (grey line), both applied on the target by the sinapinic acid double layer method. Main values of the $[M+H]^+$ peaks are shown in the figure.

Figure 1.

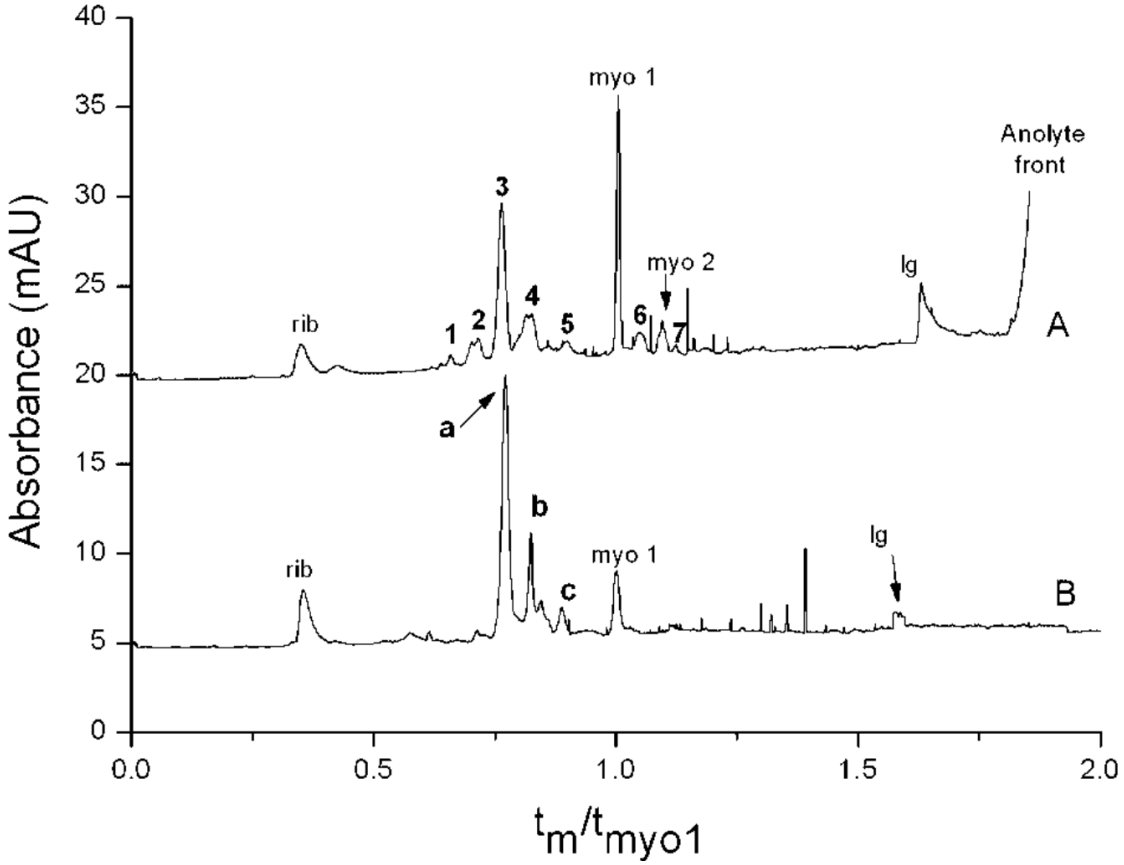
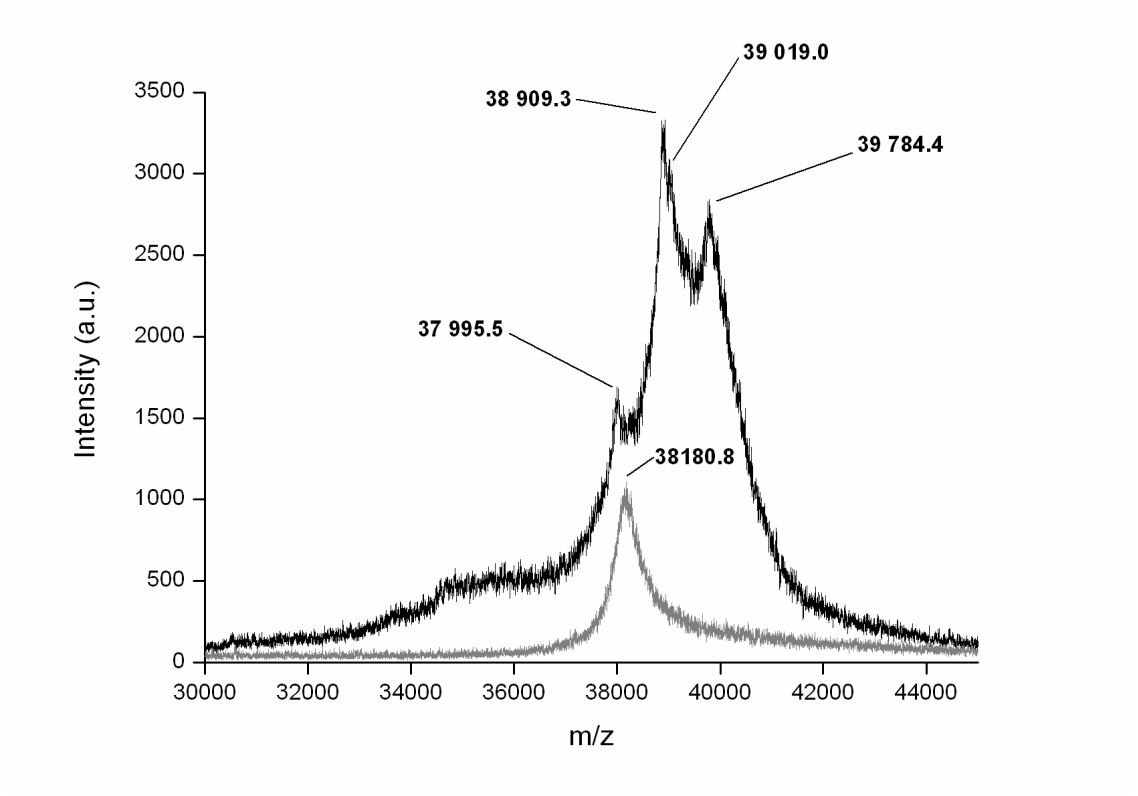

Absorbance (mAU)
40
35
30
25
20
15
10
5
0
0.0
0.5
1.0
1.5
2.0
t_m/t_{myo1}
myo 1
Anolyte front
3
myo 2
lg
4
rib
1
2
5
6
7
A
a
b
lg
myo 1
rib
c
B

Figure 2.


39 019.0
38 909.3
39 784.4
37 995.5
38180.8
Intensity (a.u.)
m/z
0
500
1000
1500
2000
2500
3000
3500
30000
32000
34000
36000
38000
40000
42000
44000