Proteomics of primary mesenchymal stem cells

S Roche1,2, M Provansal1, L Tiers3, C Jorgensen3 & S Lehmann1,2†

1Author for correspondence
2Institut de Génétique Humaine du CNRS, 141 rue dela Cardonille, 34396 Montpellier, France
3INSERM U475, 99 rue St Eloi, Laboratoire de Génétique du CNRS, 141 rue dela Cardonille, 34396 Montpellier, France
dehumaine.fr
†Author for correspondence

Keywords: 2D electrophoresis, differentiation, mass spectrometry, mesenchymal stem cell, proteomics

Tissue and functional regeneration takes place in the body at various stages throughout life. However, bone, cartilage, tendons, blood vessels and cardiac muscle have a limited capacity for self repair and, after injury or disease, the regenerative ability of these adult tissues is often insufficient and leads to nonfunctional scar tissue. In this context, mesenchymal stem cells, which are adult multipotential progenitors of mesoderm cells (osteoblasts, chondrocytes, adipocytes and stroma cells), represent a major hope for tissue-engineered replacement and regenerative medicine. Furthermore, the autologous use of these cells prevents immunological responses against new tissues and the risks of disease transmission from donors, which are both common problems of organ transplantation.

While the existence of mesenchymal stem cells is undisputed, many questions remain regarding their self-renewal and capacity to differentiate, their homogenous nature as a cell population throughout the body and their true potential in regenerative medicine. In this article, the proteomics studies carried out to characterize mesenchymal stem cells and to help understand their physiology are reviewed.

Origin & definition of mesenchymal stem cells

Bone marrow is composed of two major types of stem cells: hematopoietic stem cells and non-hematopoietic stem cells, also called mesenchymal stem cells (MSCs) or marrow stromal cells [1]. MSCs are pluripotent stem cells and are precursors of various mesoderm-type cells. They can be identified, to some extent, by their expression of various cluster of differentiation (CD) antigens [2], in particular CD 105, CD 73 and CD 90, which are considered to be classical markers for these cells [3]. In addition, they express STRO-1, a marker used to isolate multilineage progenitors from bone marrow [4]. In contrast, the hematopoietic lineage markers CD 31, CD 34 and CD 45 are not commonly observed on these cells. MSCs can be easily isolated from bone marrow, expanded in culture and differentiated in vitro into a variety of cell types, such as osteoblasts, chondrocytes, adipocytes, skeletal and smooth muscle cells [1]. Therefore, they could provide remarkable clinical potential in tissue regeneration and engineering protocols. MSCs can also be isolated from other tissues, such as cartilage [5], synovial membrane [6,7], adipose tissue [8] and umbilical cord blood [9]. However, for practical reasons (amount and availability of the tissues in humans), bone marrow, adipose tissue and umbilical cord blood are the three main sources for MSCs in regenerative medicine and will be the focus of this review.

Why study MSC proteomics?

Given the wide array of potential targets involved in the regulation of MSC self renewal and regenerative capacities, proteomics is a powerful approach that allows the detection of proteins carrying post-translational modifications, such as proteolytic maturation, glycosylation or phosphorylation, all of which are obscure in genomic studies. This is particularly relevant since signaling pathways and phosphorylation of transcription factors are most likely involved in the definition of MSC self renewal and regenerative capacities.

Proteomic technologies

High-throughput proteome analysis is a new field in basic and clinical research that is advancing rapidly towards screening of complex biological samples. The proteome can be defined as the protein constituents of a biological fluid, cell or tissue, characterized in terms of amount, post-translational modifications, interactions and turnovers. It is of an intrinsically dynamic
nature, with subtle and rapid adjustments. The sequencing of the human genome has substantially facilitated the identification of the functional products of the estimated 30,000 genes. However, numerous post-translational modifications lead to a far greater diversity in protein expression, with approximately 100 million possible functional protein variants envisioned [12]. Importantly, protein biochemical characteristics are as varied as the protein diversity, which has important consequences for proteomics sample preparation methods, in terms of solubility and extraction in buffers containing different types and amounts of detergent and denaturing agents.

2D polyacrylamide gel electrophoresis (2DE) is still the core technology for proteomics. It involves the separation of proteins from complex mixtures according to two independent parameters in two distinct steps: first, proteins are separated according to their isoelectric point on a pH gradient under denaturing conditions; second, the proteins are separated according to their apparent molecular size by SDS-polyacrylamide gel electrophoresis (Figure 1).

**Table 1. Summary of proteomics methods used for mesenchymal stem cell analysis.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Separation methods</th>
<th>Staining methods</th>
<th>Identification methods</th>
<th>Prefractionation</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2DE</strong></td>
<td>pH and molecular weight</td>
<td>Coomassie and silver staining, fluorescent or immunological detection</td>
<td>In-gel digestion, peptide mass fingerprint or de novo sequencing (MS-MS or Edman)</td>
<td>Possible: IMAC, IEX or immunopurification</td>
<td>• Core technology for proteomics</td>
<td>• Low analytical power for low abundance, hydrophobic or high-molecular-weight proteins • Time-consuming, manual technique</td>
</tr>
<tr>
<td><strong>DIGE</strong></td>
<td>pH and molecular weight</td>
<td>Covalent and fluorescent labeling prior to separation</td>
<td>In-gel digestion, peptide mass fingerprint or de novo sequencing (MS-MS or Edman)</td>
<td>Possible: IMAC, IEX or immunopurification</td>
<td>• Improvement of gel comparison</td>
<td>• Low analytical power for low abundance, hydrophobic or high-molecular-weight proteins • Time-consuming, manual technique • Cost • Labeling</td>
</tr>
<tr>
<td><strong>LC-MS</strong></td>
<td>Chromatography after tryptic digestion</td>
<td>No staining, direct detection by MS</td>
<td>Sequencing by tandem MS (LC-MS-MS)</td>
<td>Possible: for the peptides (phosphorylated)</td>
<td>• Reduction of biochemical variability of proteins • Digestion into peptides</td>
<td>• No quantification • Digestion destroys crucial information on isoforms</td>
</tr>
<tr>
<td><strong>ICAT</strong></td>
<td>Tryptic digestion and LC-MS</td>
<td>Isotope covalent labeling on cysteine residues after extraction</td>
<td>Sequencing by tandem MS (LC-MS-MS)</td>
<td>Purification of labeled peptides</td>
<td>• Reduction of biochemical variability of proteins • Digestion into peptides</td>
<td>• Only cysteine residues are labeled • Digestion destroys crucial information on isoforms</td>
</tr>
<tr>
<td><strong>SILAC</strong></td>
<td>Tryptic digestion and LC-MS</td>
<td>Labeling with light and heavy amino acid isotopes during culture</td>
<td>Sequencing by tandem MS (LC-MS-MS)</td>
<td>Possible: for the peptides (phosphorylated)</td>
<td>• All proteins and peptides are labeled</td>
<td>• Mostly restricted to cell culture • Digestion destroys crucial information on isoforms • Cost</td>
</tr>
</tbody>
</table>

2DE: 2D electrophoresis; ICAT: Isotope-coded affinity tag; IEX: Ion exchange chromatography; IMAC: Immobilized metal-ion affinity chromatography; LC: Liquid chromatography; MS: Mass spectrometry; SILAC: Stable isotope labeling by amino acids in cell culture.
Proteins in 2D E gels are often detected by silver or coomassie staining and, alternatively, by fluorescent (Sypro® Ruby, difference gel electrophoresis [DIGE] or ProQ Diamond™) or immunological methods (notably, for phosphoproteins). Probably the biggest disadvantage of 2DE is the poor analytic power for low abundance, hydrophobic or high molecular weight proteins, as noticed by Feldmann and colleagues with human umbilical cord blood MSCs (ucbMSCs) (Table 2) [13]. This may be improved using different detergents or denaturing agents [14], although to have access to low abundant proteins, additional purification steps (subcellular fractionation, chromatography and immunopurification) or increased amounts of starting material must be used.

One way to complete 2DE analysis regarding low abundant proteins and small peptides is to analyze samples by mass spectrometry (MS) methods that are often coupled to a primary separation system, such as liquid chromatography (Table 1). In general, prior to these MS techniques, samples are digested by a proteolytic enzyme, such as trypsin, converting hydrophobic and/or larger proteins into smaller sized peptides that can then be analyzed. However, to obtain quantitative proteomics data with MS, isotope-coded affinity tag (ICAT) [15] or stable isotope labeling by amino acids in cell culture (SILAC) must be employed prior to tryptic digestion (Table 1). Recently, this method was used on human MSCs to demonstrate the influence of epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) on their differentiation capacity (Table 2) [16].

Proteomics signature of MSCs
The first proteomics analysis of MSCs was performed and published in 2001 by Prockop's group [17]. Using 2DE, they compared the molecular signature of a pool of bone marrow MSCs (bmMSCs) with that of a MSC subpopulation called 'RS' that demonstrated a greater potential of proliferation and differentiation [18]. The differences observed in protein expression were linked to distinct metabolic activities between the two populations. These initial proteomics findings can be compared with more recent data obtained on bmMSCs [19] and on undifferentiated human embryonic stem (ES) cells [20–22]. Comparison between subpopulations of MSCs with various capacities of differentiation and true totipotent ES cells might help identify the proteins linked to the self-renewal property that are specific to stem cells in general. The aforementioned studies revealed that 56% of proteins identified in various MSC models are also expressed by ES cells. Interestingly, 17 proteins overexpressed in Prockop's 'RS' population of MSCs were also found to be expressed by ES cells [17]. These proteins were mainly actors in protein synthesis (initiation factor 5α, elongation factor 1-α, ribosomal protein S12 and transcription factor BTF3α) and in metabolism (glyceraldehyde 3P-dehydrogenase, transketolase enoyl-CoA, lactate dehydrogenase A, transketolase and proteasome subunit α-4). However, 55% of the proteins present in MSCs and not 'RS' cells were also found to be expressed by ES cells (e.g., T-complex protein 1-α, ribosomal protein P0 and lactate dehydrogenase). Therefore, it is difficult to be certain that the 'RS' subpopulation is closer to ES cells than the whole MSC population. In addition to these results, some new data were recently published by Wang and collaborators on human primary bmMSCs treated with transforming growth factor (TGF)-β [19]. They identified 60 proteins regulated by this growth factor [19]. For example, heat-shock protein (HSP)27 levels increased with
TGF-β1 treatment and it is noteworthy that this protein was also found to be expressed specifically in the ‘RS’ Prockop population [17].

Overall, based on these first studies, it appears that it is very difficult to identify a proteomics signature specific for MSCs. Therefore, several groups completed their investigation by studying MSCs isolated from different tissues. Thus, a comparison of bmMSCs and adipose tissue MSCs (adMSCs), in conjunction with dermal and synovial fibroblasts, was performed recently using 2DE [23]. The percentage of similarity in protein expression was approximately 58% between these cell types. However, the differences observed were still not significant enough to define a specific proteomics signature for MSCs. Several additional studies have focused on MSCs isolated from umbilical cord blood (ucb) MSCs, which represent an alternative to bmMSCs [24-26]. Following 2DE, M S and database comparison, Feldman’s group attempted to identify most proteins expressed by these ucbMSCs (Tables 1 & 2) [13]. In agreement with data from Ebbesen’s group, they confirmed that these cells express detectable amounts of vimentin [25]. Many other proteins have been identified (nph3, gelsolin, α-actinin, Type VI collagen, α-tubulin, transgelin, tropomyosin, vinculin or h-caldesmon). These proteins might be involved in MSC differentiation processes, similar to the neuropolypeptide h3 (nph3), which is known to play a role in neuroectodermal adult human stem cells [27] and differentiated cells [28]. Interestingly, tropomyosin, Type VI collagen [29] and α-tubulin are important for the interaction between hematopoietic stem cells from bone marrow and the stroma [30]. The authors also detected the expression of pyruvate kinase and albumin, which might be indicative for presence of hepatic cell precursors [31], while α-actinin, transgelin, tropomyosin, vinculin and h-caldesmon are commonly found in muscle cells and cardiomyocytes. It is noteworthy that the presence of smooth muscle proteins has already been demonstrated in bmMSCs [17], adMSCs [23] and ES cells [20-22]. Taken together, these data show that ucbMSCs already express various proteins of different MSC phenotypes. The authors suggest that a specific characteristic of stem cells is the expression, in the nondifferentiated state, of proteins from different lineages and the restriction of this expression to relevant proteins during the differentiation into a specific tissue. In conclusion, the definition of a genuine MSC signature in transcriptomic [32,33] and proteomics is clearly a challenge. Comparison between available proteomics data is one aspect of this challenge, since many different paradigms, protein extraction protocols and analytical approaches were used in the different studies. In the future, federative approaches, such as the European Genostem program [101] or the HUman PRoteome (HUPO) project [34,35] focusing on MSCs and on proteomics, respectively, may help to define a MSC signature.

<table>
<thead>
<tr>
<th>Table 2. Summary of the published mesenchymal stem cell proteomics articles.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell type</strong></td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Bone marrow MSCs</td>
</tr>
<tr>
<td>Bone marrow MSCs</td>
</tr>
<tr>
<td>Bone marrow MSCs</td>
</tr>
<tr>
<td>Bone marrow MSCs</td>
</tr>
<tr>
<td>Adipose tissue-derived stem cell</td>
</tr>
<tr>
<td>Embryo MSCs</td>
</tr>
<tr>
<td>Embryonic stem cells</td>
</tr>
<tr>
<td>Umbilical cord MSCs</td>
</tr>
</tbody>
</table>

2DE: 2D electrophoresis; LC: Liquid chromatography; MS: Mass spectrometry; MSC: Mesenchymal stem cell; SILAC: Stable isotope labeling by amino acids in cell culture.
Proteomic follow-up of MSC differentiation

In addition to the aforementioned studies on undifferentiated MSCs, an important amount of proteomics research has focused on the modification of the proteome linked to the differentiation of MSCs. In fact, these cells are multipotent and, in vitro, they maintain their ability to differentiate along several committed phenotypes including: osteoblasts, chondrocytes, adipocytes and muscle cells [1]. As for other stem cells, the commitment of MSCs to one of these phenotypes is likely driven by a variety of factors, some of them identified (bone morphogenetic protein [BMP]2, TGF-β1 or TGF-β3 for chondrocyte differentiation) and others still unknown. Using proteomics, different groups investigated the molecular mechanisms underlying the commitment process by studying the differentiation of MSCs into osteoblasts, adipocytes, muscle cells or chondrocytes [16,19,23,36–38]. Many proteins from the cytoskeleton or involved in metabolism, stress or protein synthesis had their expression modified when the undifferentiated and differentiated state of the cells were compared. For example, Lee and colleagues detected 30 protein spots on 2DE that were modified following the differentiation of bmMSCS into adipocytes [38]. Among these, they could identify, by MS, only eight proteins, which illustrated the frequent difficulties to formally identify a protein. Four of these proteins were directly associated with adipogenesis (syntaxin, oxysterol binding protein [OSBP]-related protein 3, peroxisome proliferation-activated receptor [PPAR]-γ and glycophorin). Importantly, when adMSCS have been used to generate adipocytes, the results were different [23]. In addition, data obtained with human and murine bmMSCS differentiated into adipocytes were also substantially different [23]. These discrepancies might be explained by technical differences between studies and/or by the hypothesis that MSCs from different tissue origins differ in their differentiation molecular pathways.

For differentiation into chondrocytes, Lee and colleagues have also demonstrated the induction of different extracellular matrix proteins, such as Type II collagen, matrilin, 3-phosphoadenosine-5-phosphosulfate (PAPS)-synthase and carbonic anhydrase (CA)-II, by 2DE [37]. Following this, using specific inhibitors they investigated the signaling pathways potentially responsible for this phenotype. They observed that inhibition of the ERK pathway favored the chondrocyte phenotype, while protein kinase Cα- and P38-pathway inhibition had an opposite effect. This might be linked to the activity of the Sox-9 transcription factor involved in chondrocytes differentiation [38]. Consistent with this hypothesis, the authors have observed, using 2D E, that Sox-9 is post-translationally modified during chondrocyte differentiation induced by BM P2 in the C3H10T1/2 cell line (Unpublished Data).

Finally, a recent article describes a global proteomics approach based on SILAC, phosphopeptide selection and LC-M S-M S (Tables 1 & 2) to assay post-translational protein modifications in a model of osteogenic differentiation of bmMSCS [16]. In this work, the authors demonstrated that EGF and PDGF growth factors modulate osteogenic capability of MSCs through mitogen-activated protein kinase (MAPK)/ERK, p38 kinase and phosphatidylinositol 3 kinase signaling. Exploitation of these results in order to improve and better control differentiation of MSCs into cells usable for regenerative medicine is now underway in several research groups.

Altogether, these studies on MSC differentiation were more informative than those focusing on undifferentiated cells. Moreover, they illustrate nicely the potential of proteomics studies to progress towards the identification of cellular factors and transduction pathways important for stem cell differentiation.

Conclusion

Published proteomics studies on MSCs already provided relevant information regarding the characterization of these cells and the factors associated to their differentiation. However, it is still rather difficult to generate a definitive proteomic profile of undifferentiated and differentiated MSCs from these data, owing to variation between the proteomics technologies, cell models and cell origins used in the different published studies. However, such profiling is essential to characterize MSCs and identify pathways involved in self renewal, multipotency and differentiation. Therefore, further proteomics work is needed, if possible in an integrated paradigm that includes cells from different origins and with access to phenotypic and genomic data on the cultures.

Acknowledgements

Supported by grants from the EC FP6 program ‘Genostem Adult mesenchymal stem cells engineering for connective tissue disorders. From the bench to the bedside and by the CNRS. We thank A. Goldsborough for critically reading the manuscript.
Mesenchymal stem cells

- Mesenchymal stem cells (MSCs) can differentiate into osteoblasts, chondrocytes, adipocytes, skeletal and smooth muscle.
- They represent a major hope for bone and cartilage replacement and regenerative medicine.
- For practical reasons (amounts and availability of the tissues in humans), MSCs are mostly isolated from bone marrow, adipose tissue and umbilical cord blood.

Biology of mesenchymal stem cells

- While the existence of MSCs is undisputed, many questions remain regarding their self-renewal and capacity to differentiate, and their homogenous nature as a cell population throughout the body.
- Additional basic research is also needed to decipher the molecular mechanisms of MSC differentiation and ultimately to control their fate.

Proteomics

- Proteomics allows for the detection of proteins carrying post-translational modifications, such as proteolytic maturation, glycosylation or phosphorylation.
- High-throughput proteome analysis is a new field in basic and clinical research, advancing rapidly towards screening of complex biological samples.
- 2D polyacrylamide gel electrophoresis (2DE) is still the core technology for proteomics.
- One approach to complete 2DE analysis of low abundance proteins is to analyze samples by mass spectrometry (MS).

The proteomic signature of mesenchymal stem cells

- Comparison between subpopulations of MSCs with various capacities of differentiation might help identify the proteins linked to the self-renewal property that are specific of stem cells.
- The definition of a genuine MSC signature is a challenge since many different paradigms, protein extraction protocols and analytical approaches were used in published studies.

Proteomic approaches to tracking mesenchymal stem cell differentiation

- An important amount of proteomics work has focused on the modification of the proteome linked to the differentiation of MSCs.
- Many proteins from the cytoskeleton or involved in metabolism, stress or protein synthesis have their expression modified when the undifferentiated and differentiated state of the cells are compared.
- Additional proteomic work is needed to progress towards the identification of cellular factors and transduction pathways important for stem cell differentiation.

Bibliography

Proteomics of primary mesenchymal stem cells – REVIEW


Website
101. Genostem Website
www.genostem.org