REVIEW

Clinical proteomics of the cerebrospinal fluid: Towards the discovery of new biomarkers

Stéphane Roche1, 2*, Audrey Gabelle1, 3* and Sylvain Lehmann1, 2

1 CNRS, Institut de Génétique Humaine UPR 1142, Montpellier, France
2 CHU Montpellier, Laboratoire de Biochimie, Hôpital Saint Eloi, Montpellier, France
3 CHU Montpellier, Service de Neurologie, Hôpital Gui de Chauliac, Montpellier, France

The cerebrospinal fluid (CSF) circulates within the CNS where it plays an essential physiological role in homeostasis of neuronal cells. This biological fluid has an important protein diversity that results from both filtration of serum through the blood–brain barrier and production/secretion of neuronal peptides and proteins. Changes in CSF composition depend on blood proteome, CSF circulation alterations, as well as physiological or pathological brain status. Hence, CSF proteomic analysis gives a unique opportunity to detect and describe biomarkers in neurological affections. Although lumbar puncture is considered as invasive, post lumbar puncture events remain rare and minor. Nevertheless, CSF biological analysis is currently limited to a small number of parameters and clinical situations suggesting mainly meningitis, malignancies or dementia. Few CSF proteomic studies have been performed in comparison to those on blood. In this review, we will provide a proteomics description of the CSF, summarize the current clinical use of this fluid and describe its clinical proteomics examination.

Keywords:
Biomarker / Cancer / Cerebrospinal fluid / Neurodegenerative disease

1 Cerebrospinal fluid (CSF) origin and global composition

The CSF is produced in the CNS by two different mechanisms. First of all, in the brain ventricles, the choroid plexuses which are formed by the invagination of the pia mater continuously secrete CSF constituents originating from the blood, following active and passive transports [1]. These structures are heavily vascularized and covered by modified ciliated ependymal cells linked by tight junctions materializing the so-called “blood–brain barrier”. The second origin of CSF components is represented by the drainage of the interstitial liquid of the nervous tissues (which is devoid of lymphatic system). The inflow of the interstitial liquid occurs either through the ependyma towards the ventricular liquid, or through the layer of superficial glia and the pial tissue towards subarachnoid spaces. This origin is particularly important since it explains why the CSF is a “reporter” of the physiological and the pathological status of the CNS.

The CSF is constantly reabsorbed mostly at the level of the arachnoid granulations located near the superior longitudinal venous sinus and at the radicular emergences of the spinal nerves. Hence, the CSF circulates from the ventricular crossroads to the third ventricle, through the mesencephalic aqueduct in the fourth ventricle, and reaches the subarachnoid cisterns by the median gap from the fourth ventricle (Magenie foramina). Then, the CSF flows upward around the cerebral hemispheres towards the superior sagittal venous sinus (and the blood), and also downward around the spinal subarachnoid space where it could be collected by

Correspondence: Professor Sylvain Lehmann, CNRS, Institut de Génétique Humaine UPR 1142, 141 rue de la Cardonille, Montpellier F-34396, France
E-mail: sylvain.lehmann@igh.cnrs.fr
Fax: +33-4-99619931

Abbreviations: AD, Alzheimer’s disease; CJD, Creutzfeldt Jakob disease; CSF, cerebrospinal fluid

* Both authors contributed equally to this work.
lumbar puncture. Importantly, CSF composition varies slightly with its collection site (ventricular or lumbar). In fact, CSF constituents often have a rostral-caudal gradient (i.e., concentration may be higher in ventricles or increase as CSF is removed, reflecting a greater brain contribution).

The role of the CSF is multiple. It includes a mechanical protection of the CNS against trauma and an important metabolic function through the transport of various constituents (hormones, etc.) and metabolites, ensuring the homeostasis of CNS cells.

Normal CSF contains less than 5 mononuclear cells per mL and no red blood cells. CSF glucose concentration corresponds generally to 2/3 of blood glucose. The CSF normal protein concentration is in the range of 20–40 mg per 100 mL in the lumbar cistern level [2], a value 100-times lower than in the blood. It is considered that 80% of this protein content derives from blood (through the choroid tissues) and 20% originates directly from the CNS [3]. The major proteins in both fluids are found similar except for high molecular weight proteins such as multimeric IgM that are not well transported through the blood–brain barrier. A comparison between CSF and plasma protein distribution (Fig. 1) reveals minor differences for some major blood proteins like IgG, transferrin or alpha-1-antitrypsin, while more important variations for others like albumin, IgM, or prealbumin exist. These major proteins represent 80% of the bulk protein mass in CSF, and up to 90% in serum (Fig. 2). All these specific CSF characteristics have to be taken into account for clinical proteomics procedures especially those involving the removal of major proteins [4–6].

The CSF is secreted from several different CNS structures, and any changes in the CSF composition will accurately reflect pathological processes. Proteomics offers a comprehensive bird’s eye view to analyze CSF proteins at a systems level. This paper reviews the variety of analytical methods that have been used for proteomics analysis of CSF, including sample preparation, 2-D liquid and gel electrophoresis, mass spectrometry, bioinformatics, and nongel methods. The differentially expressed CSF proteins that have been identified by proteomics methods are discussed.

![Figure 1. Comparison of blood and CSF protein distribution. The distribution of major proteins in plasma and CSF are represented in each pie chart. Values in % correspond to the fraction of the total protein content. Note that IgM are almost absent from the CSF and that the overall % of low abundant proteins is higher (17%) in this fluid than in the serum (10%).](image-url)
of the cases. It is therefore essential to have access to biomarkers for effective diagnosis at an early AD stage allowing an early therapeutic intervention. The newly proposed AD biomarkers: Tau proteins (total Tau and phosphorylated forms or P-Tau) and Aβ peptides (Aβ 1-42) have been identified through physiopathology driven research. In fact, they are present in their aggregated form in the two main neurodegenerative lesions of AD, the neurofibrillary tangles and the senile plaques. CSF concentration of Tau and P-Tau are increased in AD while Aβ levels are lower in comparison with nondemented controls [8]. The sensitivity and specificity of these makers used in combination are above 80% [9] which is valuable to optimize the positive diagnosis of AD. However, their interest is limited for the identification of other dementia (Fronto temporal dementia, Lewy Body dementia, Vascular dementia, etc.), especially in the context of mixed pathologies. Interestingly however, these biomarkers could participate in the identification of a prodromal predementia state (mild cognitive impairment or MCI) and its associated risk to develop an AD [10]. In addition to Tau and Aβ biomarkers, other candidate AD biomarkers have been proposed, based on clinical proteomics investigations (see Table 1). Some of them are promising but they need further validation to demonstrate a real clinical interest.

In the neurodegeneration field, the detection of the 14-3-3 protein by Western blot in Creutzfeldt Jakob disease (CJD) is another example of a biomarker of clinical interest (Fig. 3). The discovery of this marker is, to our knowledge, a unique and encouraging example where a pure proteomics approach, done with no a priori on the factors to identify, led to a clinical test (Fig. 3B). It followed 2-DE analyses of CJD patient CSF that led to the identification of differential protein spots as being isoforms of the 14-3-3 proteins [11–13] (Fig. 3A). The presence of this ubiquitous protein in the CSF resulted from the rapid neuronal death occurring in CJD. 14-3-3 protein detection allows the positive diagnosis of sporadic CJD with a sensitivity and specificity above 90%.

2.2 Autoimmune diseases and multiple sclerosis

Ig detection in the CSF has been used for a long time as a diagnosis criterion for multiple sclerosis which is the most common autoimmune disorder in neurology. The well-known McDonald’s diagnosis criteria for multiple sclerosis include an oligoclonal profile of CSF Igs, a test that is currently performed in clinical laboratories using isoelectrical focusing. Several clinical proteomics studies attempted to detect in multiple sclerosis, CSF specific autoantibodies. For example, autoantibodies against the myelin basic protein (a constituent of the myelin sheets damaged in multiple sclerosis) were detected in 12% of the cases in the early stage of the disease, a percentage that increased during multiple sclerosis progression [14]. It is possible that multiple detection of myelin autoantibodies, even in the blood [15], could be useful for the biological follow up and the detection of multiple sclerosis. However, this remains to be confirmed.
Clinical proteomics approaches looking at the presence of autoantibodies are not restricted to autoimmune diseases and are now often used in cancer research. Interestingly, we recently adapted to the CSF the use of protein chips for the screening of autoantibody (Fig. 4). We expect that this clinical proteomic approach might lead to the identification of new biomarkers in various neurological disorders.

3 Proteomics analysis of the normal CSF proteome

Most clinical proteomics studies of the CSF were conducted to identify new biomarkers in different disorders. Hence, only a few studies focused on the global CSF proteomics composition which is as complex as that from the blood. Importantly, in addition to the CSF proteins originating from the blood and the CNS, some constituents could also come from cells present in the fluid itself (macrophages, lymphocytes, bacteria, etc.). In the normal situation however, this protein origin is minor.

CSF composition varies also with the position of the puncture (lumbar or ventricular puncture) and, hopefully with the physiopathological status of the patient.

CSF proteins can be sorted by their relative amount (Fig. 1) or regrouped according to their physiological function. Many proteins with similar function could be found in CSF and serum [16]: for example, carrier specific and non-specific proteins (transferrin, haptoglobin, hemopexin, or albumin), enzymes (prostaglandin D2-synthase, alpha-2-
4 Rationale and specificities of CSF clinical proteomics investigations

The development of neuroscience allowed for a better understanding of the physiopathology of many neurological disorders. It also contributed to the optimization of diagnostic criteria (see above paragraphs), improvement of an early care, and the development of innovating therapeutics aiming at blocking the affection process. However, despite these advances, diagnosis, prognosis and follow up of neurological disorders remain a major challenge. Biomarkers that allow a quantitative and multisite way to characterize/follow up a pathological situation are therefore well needed in neurology. Most discovery programs are focusing on the CSF as this fluid is in direct contact with the extra-cellular space of the CNS.

macroglobulin, alpha-1-antitrypsin, ceruloplasmin, lysosome, enolase) or structural subunit (fibrinogen, GFAP, myelin basic protein). Immune response and inflammatory proteins are also well represented (Igs, complements C3 and C4, C-reactive protein, serum amyloid A, orosomucoid). Some protein constituents come from the brain (cystatin C, glial fibrillary acid protein, myelin basic protein, S100) or have been modified in the CNS (beta-2-transferin, fibrinogen degradation products). Importantly, due to ethical consideration and sample availability, most of the CSF clinical proteomics have been using “control” samples such as normal pressure hydrocephaly.

At the analytical level, 2-D PAGE (2-DE) was still the core technology used for CSF proteomics studies [17–19]. Two recent papers provided a comprehensive view of the global protein composition of the CSF [4, 5]. Interestingly, in one of this work, the removal of high-abundant proteins by immunocapture and ultrafiltration allowed the detection by 2-D LC-MS/MS analysis of several new sets of unique CSF proteins. The study of the global CSF proteome was also completed by peptidomics investigations which revealed the presence in this fluid of many neuroendocrine peptides and neurosecretory proteins [20–22].

Finally, an exhaustive view of CSF proteome was obtained by Fourier transform ICR which could identify more than 4000 different proteins [23, 24].
This fluid could therefore be considered as “a window to the brain” giving access to its physiopathological status. In fact, biological changes in the brain could induce modifications in the CSF biochemical composition due to the release of cellular components, cytokines, peptides, hormones, etc. The necessity to perform a lumbar puncture to collect the CSF represents one of the major limitations of the systematic analysis of this fluid. It raises ethical questions (especially for healthy controls), even though lumbar puncture is more and more considered as a safe and acceptable procedure. In fact, it is associated with a very low risk of infection and the apprehended post lumbar puncture headaches are rare, especially among older patients [25]. Incidentally, the procedure using flexible needles and cutaneous or gas anesthesia could limit the lumbar puncture anxiety and the possible pain.

One additional hurdle of CSF clinical proteomics studies resides in the difficulties to realize accurate diagnosis in neurology and to compare matched groups of patients and controls. This is particularly true as neurology patients sometimes have mixed disorders and/or associations of different and unrelated pathologies, in particular in elderly people. If possible anatomopathological data need to be obtained as it often is the only way to have confirmed diagnosis.

As for serum and plasma [26], clinical proteomics analyses of the CSF need to pay a specific attention to preanalytics. This is especially true in the context of rare neurological disorders that necessitate the involvement of several clinical centers to constitute large cohorts. Delays and temperature between collection and processing, adapted centrifugation and storage procedures are among the important steps to be controlled. The importance of preanalytics for CSF biomarkers is particularly well illustrated by the Tau and Aβ markers which are temperature sensitive and can be accurately measured only if the CSF has been collected directly in polypropylene tubes [27]. To realize a valid differential clinical proteomics study, it is therefore essential that all the samples are subjected to a similar preanalytical treatment. In this context, our laboratory was appointed by the French National Institute against Cancer (INCa) to define a standard CSF preanalytical protocol. Our recommendations include drawing CSF directly into polypropylene tubes [27]. To realize a valid differential clinical proteomics study, it is therefore essential that all the samples are subjected to a similar preanalytical treatment. In this context, our laboratory was appointed by the French National Institute against Cancer (INCa) to define a standard CSF preanalytical protocol. Our recommendations include drawing CSF directly into polypropylene tubes, sample transport on ice in less than 2–4 h, centrifugation at 1000 g (10 min, 4°C) and aliquoting in low protein binding tubes before storage at −80°C.

5 Clinical proteomics investigation for new CSF biomarkers

Neurological disorders and in particular dementia and CNS metastasis, represent a growing concern in healthcare. Incidentally, we observed an increasing number of CSF clinical proteomics studies to unravel new biomarkers.

As for the blood, due to the high dynamic range and the presence of major proteins (Fig. 1), the proteomics analysis of low abundant proteins in CSF is a real challenge. One way to go around this problem is to remove CSF most abundant proteins by immunoaffinity columns [4, 6] as done in blood [28]. However, as the distribution of the major proteins is different in these two fluids, the efficiency and the limit of this approach has to be carefully assessed (Fig. 2). For example, Prostaglandin D2-synthase, Apolipoprotein E or Cystatin C, that are major proteins of the CSF, are not removed by immunodepletion kits designed for the blood. In addition, the interest of depletion depends on the analytical technologies: immunocapture [6] and hydrophobic affinity chromatography [21] for 2-DE, size fractionation for peptidomics [29]. One important issue for clinical proteomics is the reproducibility and the robustness of these prefractionation methods.

Many biomarkers have been proposed in neurological disorders following classical 2-DE, LC-MS/MS, FT-ICR or profiling technologies combining chromatography on chip (SELDI-TOF) or on magnetic microbeads (Clinprot®) followed by a TOF profile as discriminating technology (Table 1). It is difficult to be exhaustive on this subject. Below, we report some examples of clinical proteomics studies in dementia and cancer research.

Recently, a 2-DE differential analysis performed on the CSF of 34 AD patients and 34 age-matched controls identified 23 differential spots. They corresponded to proteins involved in Aβ transport, inflammatory response, proteolytic inhibition, and neuronal cell function. The authors suggested that this panel of proteins could be used to differentiate AD and non-AD gels. In a blind study, a sensitivity of 93% and a specificity of 90% for AD diagnosis [29] was obtained with this panel. Another study identified cystatin C as a potential biomarker for Alzheimer disease with a very high specificity [30]. However, 2 years later the same group published a correction indicating that this result was an artifact linked to the storage of the samples [31] which is a good example of the importance of preanalytical phases.

For the diagnosis of CNS tumors, CSF cytological examination often gives low (positive findings in less than 50% of the cases) and nonspecific results [32]. Based on elevated serum levels of some proteins (carcinoembryonic antigen, β-2 microglobulin, α-feto protein, etc.) in systemic malignancy, CSF clinical proteomics studies initially focused on similar subset of markers. However, this approach was not helpful for diagnosis or neoplastic classification [33]. Recent studies with SELDI-TOF have identified protein profiles in the blood that could discriminate malignant and nonmalignant situations. Using a similar approach in the CSF, several biomarker patterns have recently been proposed to identify patients with neoplastic and reactive/inflammatory CNS diseases [34]. With 2-DE and cleavable Isotope-coded affinity Tag (iCAT), a recent study also identified 20 biomarkers that could serve as new diagnostic, prognostic and disease follow-up markers when used alone or in combination [35].
6 Conclusion

CSF clinical proteomics represents a major tool for biomarker discovery in the neurological field. Neurological disorders are in fact difficult to diagnose and represent a growing concern for public health with the increase of AD and malignant CNS affections. CSF biomarkers appear therefore as invaluable tools in terms of diagnosis, prognosis of the course of CNS disease, and as a predictive value in the presymptomatic state. However, CSF proteomics analysis is not trivial, as this fluid combines hurdles of blood analysis (complexity, dynamic range, etc.) and specific issues in terms of collection and physiological variations. So far, few studies resulted in significant clinical application, but the use of prefractionation methods and new high-end proteomics technologies on CSF collected in controlled preanalytical conditions could lead to future relevant biomarkers.

Supported by grants from the EC FP6 Program cNEUPRO, the Institut National Contre le Cancer (INCa) and the CNRS.

The authors have declared no conflict of interest.

7 References


© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim www.clinical.proteomics-journal.com


