Identification of novel candidate protein biomarkers for the post-polio syndrome — Implications for diagnosis, neurodegeneration and neuroinflammation

Henrik Gonzalez\textsuperscript{a,1}, Jan Ottervald\textsuperscript{b,f,*1}, Kerstin C. Nilsson\textsuperscript{c}, Niclas Sjögren\textsuperscript{d}, Tasso Miliotis\textsuperscript{e}, Helena Von Bahr\textsuperscript{f}, Mohsen Khademif, Bodil Eriksson\textsuperscript{g}, Sven Kjellström\textsuperscript{h}, Akos Vegvari\textsuperscript{h}, Robert Harris\textsuperscript{f}, György Marko-Varga\textsuperscript{h},Kristian Borg\textsuperscript{a}, Johan Nilsson\textsuperscript{i}, Thomas Laurell\textsuperscript{i}, Tomas Olsson\textsuperscript{f,1}, Bo Franzén\textsuperscript{b,1}

\textsuperscript{a}Division of Rehabilitation Medicine, Department of Clinical Sciences, Danderyd Hospital, Karolinska Institute, Stockholm, Sweden
\textsuperscript{b}Molecular Pharmacology, AstraZeneca R&D Södertälje, Sweden
\textsuperscript{c}Disease Biology, Local Discovery, AstraZeneca R&D Södertälje, Sweden
\textsuperscript{d}Biostatistics, AstraZeneca R&D Södertälje, Sweden
\textsuperscript{e}Bioscience, Local Discovery, AstraZeneca R&D Mölndal, Sweden
\textsuperscript{f}Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Hospital, Stockholm, Sweden
\textsuperscript{g}BioPR&D, AstraZeneca R&D Södertälje, Sweden
\textsuperscript{h}Biological Sciences, Local Discovery, AstraZeneca R&D Lund, Sweden
\textsuperscript{i}Lund Technical University, Lund, Sweden

\textbf{ARTICLE DATA}

\textbf{A R T I C L E  D A T A}

\begin{tabular}{|l|}
\hline
Article history: \\
Received 1 October 2008 \\
Accepted 14 November 2008 \\
\hline
\end{tabular}

\textbf{Keywords:}
Post-polio syndrome
Pathophysiology
CSF
Proteomics
Biomarkers
Diagnosis

\textbf{A B S T R A C T}

Survivors of poliomyelitis often develop increased or new symptoms decades after the acute infection, a condition known as post-polio syndrome (PPS). The condition affects 20–60% of previous polio patients, making it one of the most common causes of neurological deficits worldwide. The underlying pathogenesis is not fully understood and accurate diagnosis is not feasible. Herein we investigated whether it was possible to identify proteomic profile aberrations in the cerebrospinal fluid (CSF) of PPS patients.

CSF from 15 patients with well-defined PPS were analyzed for protein expression profiles. The results were compared to data obtained from nine healthy controls and 34 patients with other non-inflammatory diseases which served as negative controls. In addition, 17 samples from persons with secondary progressive multiple sclerosis (SPMS) were added as relevant age-matched references for the PPS samples.

The CSF of persons with PPS displayed a disease-specific and highly predictive ($p=0.0017$) differential expression of five distinct proteins: gelsolin, hemopexin, peptidylglycine alpha-amidating monoxygenase, glutathione synthetase and kallikrein 6, respectively, in comparison with the control groups. An independent ELISA confirmed the increase of kallikrein 6.

Abbreviations: PPS, Post-polio syndrome; CSF, cerebrospinal fluid; SPMS, secondary progressive multiple sclerosis; HC, healthy controls; OND, other neurological disease; PLS, Partial Least Squares; VIP, Variable Importance in the Projection; PAM, peptidylglycine alpha-amidating monoxygenase; Mr/pI, molecular weight/isoelectric point.

* Corresponding author. Neuroimmunology Unit, Centre for Molecular Medicine, CMM, L8:04, Karolinska Hospital, 171 76, Stockholm, Sweden. Tel.: +46 8 51776246, +46 73 6459165; fax: +46 8 51776248.
E-mail address: jan.ottervald@ki.se (J. Ottervald).

\textsuperscript{1} Equal contribution.

1874-3919/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jprot.2008.11.014
1. Introduction

The post-polio syndrome (PPS) appears later in life after an acute poliomyelitis infection with a prevalence of 20–60%. Approximately twenty million persons worldwide are affected to some degree. This makes PPS one of the most prevalent motor neuron diseases [1]. Survivors of acute paralytic disease later in life develop symptoms such as increased weakness of the musculoskeletal system, fatigue, pain, gait disturbances, breathing difficulties as well as swallowing problems [2]. Consequently, it may be difficult to differentiate symptoms of an active PPS disease from those of normal aging. The current diagnosis is based on thorough clinical examinations in order to eliminate other possible causes, and thus identification of diagnostic biomarkers is highly desirable. Furthermore, identification of any protein abnormalities may give insight of the underlying pathophysiology, which in turn could indicate new therapeutic options.

An ongoing denervation has been suggested as the cause of the increased weakness in muscles earlier affected by poliomyelitis [3,4]. This denervation is compensated by reinnervation by means of collateral sprouting, leading to an increase of the area of the motor units. However, reinnervation cannot sufficiently compensate for denervation, leading to whole or partial loss of motor units followed by a decrease of muscle strength [4]. It is debated whether this deterioration is due to further loss of motor neurons due to normal aging, deleterious over-use of remaining motor neurons, and/or an active disease process, perhaps involving chronic intrathecal inflammatory damage [5–19].

To study the pathophysiology of PPS, CSF is of particular interest since its protein content is expected to reflect changes in the target organ. Previous studies of CSF proteins have suggested associated biomarkers in the neuroscience fields of Creutzfeldts disease [20] and Alzheimer’s disease [21,22]. In the present study we applied classical proteomics to take an unbiased approach to study a broad spectrum of proteins in the CSF from PPS patients and control subjects with or without other central nervous system disease. The aim was to identify candidate biomarkers for PPS and further to uncover hitherto unknown biological processes related to PPS.

2. Methods

2.1. Patient population

Consecutive PPS patients with a history of acute poliomyelitis and new problems such as increased muscle weakness, muscle fatigue and pain in muscle groups earlier affected by poliomyelitis according to the criteria for PPS as given by Halstead and Rossi [23] were included. (Supplementary data 1.1) All 15 PPS patients in this study had been examined before with regard to CSF mononuclear expression of TNF-α and IFN-γ and all displayed elevated mRNA expression in CSF of these cytokines as compared to in controls [24] (Table 1). Control patient CSF samples were either from non-disease volunteers, healthy controls (HC), or persons with neurological signs or symptoms for other reasons than PPS, ‘other neurological disease’ (OND), respectively. In the following text both healthy controls and OND patients are grouped together as ‘OND/HC’ n = 43. These patients were regarded as non-inflammatory and the lumbar puncture had been performed as a part of routine clinical analysis of these patients. All samples were assessed for erythrocyte contamination before analysis. The CSF of these patients displayed no signs of inflammation in the form of pleocytosis, increased IgG index or oligoclonal bands. A comparison with samples from persons with secondary progressive Multiple Sclerosis (SPMS) was added (n = 17). SPMS samples were age-matched to the PPS group and can be considered to be an inflammatory disease reference material (Table 1). CSF from an independent population of persons with PPS and OND were used to validate Kallikrein 6 expression levels. CSF was obtained from 37 individuals with PPS and 30 individuals diagnosed as being OND (Table 1).

2.2. Ethics

All study enrolment followed the recommendations of the declaration of Helsinki and the study was approved by the Ethics committee of the Karolinska Institute #02-365 and #02-213. Oral and written information was given to the patients and confirmed consent in writing was received before inclusion.

Table 1 – The diagnosis and number of patients used in the study

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Number</th>
<th>Male</th>
<th>Female</th>
<th>Mean age male</th>
<th>Mean age female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy controls*</td>
<td>9</td>
<td>0</td>
<td>10</td>
<td>–</td>
<td>32.4</td>
</tr>
<tr>
<td>OND*</td>
<td>34</td>
<td>12</td>
<td>22</td>
<td>42.1</td>
<td>38.6</td>
</tr>
<tr>
<td>Non-inflammatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPMS*</td>
<td>17</td>
<td>6</td>
<td>11</td>
<td>54.1</td>
<td>56.5</td>
</tr>
<tr>
<td>Inflammatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPS*</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>59.3</td>
<td>60.5</td>
</tr>
<tr>
<td>OND ELISA evaluation</td>
<td>30</td>
<td>9</td>
<td>21</td>
<td>37.7</td>
<td>43</td>
</tr>
<tr>
<td>Non-inflammatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPS ELISA evaluation</td>
<td>37</td>
<td>16</td>
<td>21</td>
<td>61.5</td>
<td>56.9</td>
</tr>
<tr>
<td>Total number of patients</td>
<td>n=142</td>
<td>52</td>
<td>90</td>
<td>Total mean 50.9</td>
<td>Total mean 47.9</td>
</tr>
</tbody>
</table>

Patients included in the proteomic part indicated with *.
2.3. Sample preparation and protein separation

All samples were albumin- and IgG-depleted to increase the sensitivity and resolution of the analyses. Preparation and protein concentration prior to 2-DE is described in Supplementary data 2. Samples was analyzed by 2-DE (isoelectric focusing followed by SDS-PAGE) as described elsewhere [25], thereafter stained with fluorescent dye [26] and finally digitized with a laser scanner. All samples were analyzed in a randomized sequence to avoid batch-to-batch influence of the results. Image analysis and quantitative protein detection were made prior to analysis using PDQuest™ software (version 7.3, Bio-Rad, Hercules, CA, USA). The Mr/pl coordinates were calculated from previously known and identified proteins from CSF (http://expasy.org/swiss-2dpage/viewer.)

2.4. Data analysis

Image analysis of the 2DE gels resulted in 1,499 detected and matched protein spots. All data was pre-processed and analyzed as described in Supplementary data 3.

With an aim to identify a reduced number of protein spots with predictive value we applied multivariate predictive modeling to the data. The modeling was carried out in two steps. First, a training set consisting of 36 OND/HC and nine PPS was used to select the protein spots with the highest predictivity and to optimize model parameters. In the next step the predictivity of the model and the associated selected protein spots was evaluated using an independent test set consisting of six PPS and seven OND/HC samples (Fig. 1).

Additionally, we compared the differentially expressed proteins in CSF from PPS with expression levels of same proteins in CSF from SPMS (17 age-matched patients).

We used the Partial Least Squares (PLS) method [27] to build predictive models of the data. To optimize model parameters and to select important protein spots we applied a cross-validation procedure to the training set data, coupled with a variable (corresponding to a protein spot) selection procedure. Briefly, in the cross-validation procedure the training set is internally divided into training and test sets randomly, and this is repeated a number of times. The prediction rate is recorded as an average of the test set predictions over all cross-validation rounds. Note that the variable (protein spot) selection is only performed on the training set, as inclusion of the test set in the variable selection can result in substantial overestimation of the predictive ability of the model [28]. The model optimization was performed on two datasets resulting from two different approaches to handling missing values (Supplementary data 3.1).

Finally, a test set evaluation was then performed to assess the predictivity of the selected model and associated protein spots using six PPS and seven OND/HC independent samples.

The PLS modeling was performed using Matlab (The Mathworks, Inc., Natick, MA, USA) and the PLS toolbox (Eigenvector Research, Wenatchee, WA, USA). Hierarchical clustering was performed using Spotfire (http://www.spotfire.com).

2.5. Protein identification

Protein spots of interest were excised from gels using the EXQUEST spot cutter robot (Bio-Rad, Hercules, CA, USA) and transfered to 96-well plates. Up to 6 protein spots of the same spot identity were pooled to each well in order to facilitate the identification of low abundant protein spots. The excised gel plugs were subjected to destaining using a wash solution consisting of 70% ACN (LC.MS grade acetonitrile, Fischer Scientific, Loughborough, UK) in 25 mM ammonium bicarbonate (Sigma, St. Louis, MO, USA) for 10 min. This washing procedure was repeated three times. Finally, gel plugs were speed vacuum dried until dryness. Reduction was performed at 56 °C for 30 min by addition of 20 µL of 20 mM DTT solution (Genomic Solutions Inc., Ann Arbor, MI, USA). After cooling of samples to room temperature 20 µL of 55 mM iodoacetamide were added and samples were alkylated for 45 min in the dark. Supernatants were removed and ACN was added to shrink the gel plugs. About 15 µL trypsin (modified/sequencing grade, Promega, Madison, WI, USA) solution (0.075 µg) was added to the gel plugs and digestion was performed overnight at 37 °C. Extraction was performed for 1 h by adding 20 µL of extraction solution (1% ACN, Fischer Scientific, UK) 0.1% TFA, Fluka puriss. p.a. for HPLC, Steinheim, Switzerland).

During the peptide preparation two different systems were used to separate peptides onto MALDI target plates: a micro-technology workstation [29] and a nano-LC on an 1100 Nanoflow Proteomics Solution system (Agilent Technologies, Böblingen, Germany). (Supplementary data 4.1). Fractions spotted onto the MALDI targets were analyzed in MS and MS/MS mode using the 4700 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA) MALDI-TOF/TOF instrument. All MS/MS data from the MALDI-TOF/TOF instrument were acquired using 4700 Explorer software (Applied Biosystems, Foster City, CA, USA) allowing non-redundant and fully automated selection of precursors for MS/MS
acquisition (Supplementary data 4.2). Resulting MS/MS data obtained from the MALDI-TOF/TOF instrument were analyzed using GPS Explorer (Applied Biosystems, Foster City, CA, USA) software, which invoked a MASCOT 2.2.0.3 (Matrix Science, London, UK) database search using the human subset of in-house protein sequence databases (Genseq P, RefSeqP, FDB, PIR, SwissProt and TREMBL and GChuman). Mass tolerance for the precursor ion was set to 50 ppm; mass tolerance for the fragment ions was 0.2 Da. Carbamidomethylation was used as fixed modification and Oxidation (M) was allowed as a variable modification and mis-cleaveage was set to 1. MASCOT peptide hits with a MOWSE score equal or higher than 35 being considered as positive hits. A protein identity comprised of at least two tryptic fragments and/or combined with MALDI spectra scoring.

2.6. ELISA

Determinations of kallikrein 6 levels were assessed in the CSF of PPS (n=37) and OND (n=30) patients using a commercial kallikrein 6 ELISA kit (IBEX Pharmaceuticals, Inc. Montreal, Canada) according to the manufactures instructions. CSF samples were diluted 1:50 and kallikrein 6 values were calculated from a standard curve and expressed as ng/ml. Differences in CSF Kallikrein 6 levels were tested for significance using the non-parametric Mann Whitney test (GraphPad Prism 3.0, La Jolla, CA, USA).

3. Results

To date, no proteomic study has to our knowledge been performed on CSF from PPS patients. In our study there were 142 patients and controls in total, in which comparative CSF protein expression data were obtained from 15 PPS patients and 60 non-PPS controls. Moreover, protein Kallikrein 6 levels were further evaluated in 37 PPS patients and 30 OND patients with ELISA.

3.1. Model optimization and variable selection

The PLS models with one component gave the best discrimination between PPS and controls for the training set. The prediction rate remained relatively constant as variables were dropped and it was observed that about 5–15 variables were sufficient to achieve a high prediction rate of PPS (Supplementary data Fig. 1). The average test set prediction rate in the cross-validation procedure was 90% and 99% correctly predicted samples from the two datasets, respectively. With the PLS method it is possible to determine the importance of a variable (protein spot) for the model from the VIP parameter (Variable Importance in the Projection), see Supplementary data 3. VIP scores from the two models were combined to form an overall variable (protein spot) importance rank and the top 13 protein spots were chosen for the final model.

3.2. Prediction of independent test samples

To evaluate the model and the associated 13 protein spots we used an independent test set consisting of six PPS and seven OND/HC samples. The prediction result of the test set is presented in Fig. 2. Six-out-of-seven OND/HC samples were correctly predicted and all six PPS samples were correctly predicted. Thus the number of correctly predicted samples is 12/13 (92.3%), with a p-value of 0.0017 that these results would have been obtained by chance (Supplementary data 3.4).

Fig. 2–(A) Prediction of independent patient samples, the test set, using the 13-variable prediction model. The cut off-value, represented by the dashed line at 0.425, was determined from the two models originated from the two approaches to handling missing values. The number of correctly predicted samples is 12 of 13 (92.3%), with a p-value of 0.0017 that these results would have been obtained by chance (see supplementary data for details regarding the statistical test). For reference, predicted y-values of the samples used to build the model, the training set, are shown in (B).
Table 2 – Identification of the top-13 most predictive protein spots excised from 2DE gels

<table>
<thead>
<tr>
<th>Spot number (SSP)</th>
<th>Protein identity * indicates protein fragment</th>
<th>Acc.# Mascot score</th>
<th>Matched peptides</th>
<th>Coverage (%)</th>
<th>Theoretical MW (kDa)</th>
<th>Theoretical pl</th>
<th>Observed MW (kDa)</th>
<th>Observed pl (gel)</th>
<th>First amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3403</td>
<td>Gelsolin*</td>
<td>P06396</td>
<td>166</td>
<td>4</td>
<td>80,876</td>
<td>5.58</td>
<td>54</td>
<td>5.08</td>
<td>EGQQTAPASTR AGALNSDAAFLK</td>
</tr>
<tr>
<td>3408</td>
<td>Gelsolin*</td>
<td>P06396</td>
<td>192</td>
<td>4</td>
<td>80,876</td>
<td>5.58</td>
<td>54</td>
<td>5.14</td>
<td>AGALNSDAAFLK YETPDANRDR</td>
</tr>
<tr>
<td>4414</td>
<td>Peptidylglycine alpha-Amidating Monoxygenase(PAM)</td>
<td>P19021</td>
<td>186</td>
<td>3</td>
<td>97,051</td>
<td>6.00</td>
<td>51</td>
<td>5.40</td>
<td>YETPDANRDR RPTTVVK</td>
</tr>
<tr>
<td>5617</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>235</td>
<td>5</td>
<td>52,254</td>
<td>6.57</td>
<td>77</td>
<td>5.75</td>
<td>EGGQTAPASTR AGALNSDAAFLK</td>
</tr>
<tr>
<td>5631</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>259</td>
<td>4</td>
<td>52,254</td>
<td>6.57</td>
<td>77</td>
<td>5.65</td>
<td>YETPDANRDR RPTTVVK</td>
</tr>
<tr>
<td>5632</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>195</td>
<td>4</td>
<td>52,254</td>
<td>6.57</td>
<td>77</td>
<td>5.56</td>
<td>YETPDANRDR RPTTVVK</td>
</tr>
<tr>
<td>6605</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>248</td>
<td>5</td>
<td>52,254</td>
<td>6.57</td>
<td>77</td>
<td>5.84</td>
<td>YETPDANRDR RPTTVVK</td>
</tr>
<tr>
<td>6611</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>179</td>
<td>5</td>
<td>52,254</td>
<td>6.57</td>
<td>77</td>
<td>5.95</td>
<td>YETPDANRDR RPTTVVK</td>
</tr>
<tr>
<td>7208</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>34</td>
<td>1</td>
<td>52,254</td>
<td>6.57</td>
<td>34</td>
<td>6.20</td>
<td>QQHNSVFLIK</td>
</tr>
<tr>
<td>7310</td>
<td>Glutathione Synthetase(GSHB) *</td>
<td>P48637</td>
<td>54</td>
<td>2</td>
<td>52,523</td>
<td>5.67</td>
<td>38</td>
<td>6.16</td>
<td>QQHNSVFLIK</td>
</tr>
<tr>
<td>7328</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>45</td>
<td>1</td>
<td>52,254</td>
<td>6.57</td>
<td>34</td>
<td>6.05</td>
<td>QQHNSVFLIK</td>
</tr>
<tr>
<td>8206</td>
<td>Hemopexin</td>
<td>P02790</td>
<td>46</td>
<td>1</td>
<td>52,254</td>
<td>6.57</td>
<td>28</td>
<td>6.37</td>
<td>QQHNSVFLIK</td>
</tr>
<tr>
<td>8221</td>
<td>Kallikrein 6</td>
<td>Q92876</td>
<td>138</td>
<td>10</td>
<td>27,523</td>
<td>7.15</td>
<td>28</td>
<td>6.31</td>
<td>QQHNSVFLIK</td>
</tr>
</tbody>
</table>

The SSP spot numbers and observed Mr/pl refer to the 2DE gel positions shown in Fig. 2. Identification was performed by mass spectrometry as described in Section 2.5.
3.3. Identity and expression levels of proteins

Protein spots of interest were cut out from 2D-gels and identified by mass spectrometry (Table 2). A representative 2D-gel is depicted in Fig. 3 and the locations of the top ranked 13 protein spots are magnified. Each spot is given a unique database SSP number by the PDQuest™ software. Table 2 and Fig. 3 present protein identities obtained by mass spectrometry analyses together with theoretical and observed molecular weight/isoelectric points (Mr/pI). The identities were gelsolin fragments (SSP 3403 and 3408), full-length hemopexin (SSP 5617, 5631, 5632, 6605 and 6611), hemopexin fragments (SSP 7208, 7328 and 8206), kallikrein 6 (SSP 8221), peptidylglycine alpha-amidating mono-oxygenase (PAM, SSP 4414) and glutathione synthetase (GSHB, SSP 7310), respectively.

The relative importance of the protein spots for the prediction model was determined by calculating the PLS VIP scores, both for the training set and for a model consisting of all samples, i.e. the combined training and test sets. The combined model was constructed so as to use as much of the information available in determining protein spot importance (Supplementary data Table 1). Note that 11 of the top ranked protein spots represent only three different protein entities, gelsolin, kallikrein-6 and hemopexin, respectively. This indicates per se that these proteins may have a key role in PPS.

The expression levels and descriptive statistics for the top-13 proteins are presented in Table 2a Supplementary data. To provide a summary overview of the expression levels of these 11 protein spots, three meta-variables were derived based on the sum of spots with the same identity corresponding to the expression level for the proteins. Fig. 4 reveals that gelsolin fragments, hemopexin full-length and fragments and kallikrein-6 are all differentially expressed, not only compared to the control group but also compared to the SPMS group. (Expression levels and descriptive statistics are listed in Supplementary data Table 2a and b for protein spots and the derived meta-variables, respectively).

We observed a 2.4-fold increase of two protein spots representing gelsolin fragments (Fig. 4 and Supplementary data Table 2b), while no significant change in the level of full-length gelsolin was observed (data not included). This most likely reflects an increased enzymatic cleavage of the protein and possibly, in parallel, an increased compensatory production of gelsolin. Gelsolin cleavage is tightly associated with apoptosis, and both pro-apoptotic and anti-apoptotic roles have been described [30].

There was a 2.3-fold decrease of full-length hemopexin variants in parallel with a 4.2-fold increase of fragmented hemopexin, indicating a proteolytic cleavage of the protein (Fig. 4) and (Supplementary data Table 2b). Interestingly, this is in contrast to the classical role of hemopexin which belongs to the family of acute-phase proteins that are usually induced during inflammatory events [31].

The kallikrein 6 protein spot intensity increased 5.3-fold (Fig. 4 and Supplementary data Table 2a) and is likely to be a full-length variant (according to observed versus theoretical...
Fig. 4 – Box-plots showing expression levels in ppm for 43 OND/HC, 15 PPS and 17 MS-SP samples. The boxes are defined by the 25th and 75th percentile, thus 50% of the data is represented within each box. The lower/upper whiskers indicate the smallest/largest value greater/less than or equal to the first/third quartile minus/plus 1.5*interquartile range. The median is represented by a left-arrow and the mean is represented by a right-arrow.

Please cite this article as: Gonzalez H, et al, Identification of novel candidate protein biomarkers for the post-polio syndrome — Implications for diagnosis, neurodegeneration and neuroinflammation, J Prot (2009), doi:10.1016/j.jprot.2008.11.014
molecular weight) of the protein (Table 2). Kallikrein 6 is a serine protease and is preferentially expressed in the CNS at highest levels in the spinal cord and brainstem [32], and has a trypsin-like broad substrate enzymatic activity [33,34]. In addition to this single kallikrein 6 spot (SSP 8221) we observed one other spot (SSP 8209) identified as kallikrein 6 (data not included), although at lower VIP scores (VIP rank 35). Interestingly both kallikrein 6 spots exhibited increased levels in PPS, which support a possible role of kallikrein 6 in PPS. Moreover, the expression levels in the SPMS group for the proteins are at the same level as for the OND/HC-group (Fig. 4). The increased expression level of kallikrein 6 was confirmed using a commercial ELISA (methods) with a p value of 0.0007 (Fig. 5). It is not surprising to observe a difference in the fold change value of kallikrein 6 expression levels between the two methods since proteomics measure specific forms of the protein whereas ELISA assesses all forms recognized by the antibody.

Two additional proteins of possible relevance to PPS were identified: peptidylglycine alpha-amidating monoxygenase (PAM), with 3.2-fold decrease, which plays a key role in the biosynthesis of many neuronal peptides and is present in neurons as well as in oligodendrocytes [35]. It has been speculated that an increase of PAM activity in CSF from patients with MS may arise from demyelination and oligodendrocyte destruction [36]. Glutathione synthetase (GSHB), with 1.9-fold increase, is a key enzyme in the biosynthesis of glutathione. Glutathione is important for a variety of biological functions including protection of cells from oxidative damage by free radicals, detoxification of xenobiotics and membrane transport [37]. However, these two proteins had the lowest VIP scores and are thus of least significance for the predictive model.

To provide an overview of all samples in relation to the expression levels of the protein spots used in the model building, a two-dimensional hierarchical clustering was conducted (Fig. 6). Hierarchical clustering is an unsupervised method which allows data to self-organize according to similarity with respect to expression pattern. A clear separation of the two patient groups PPS and OND/HC was observed, and all samples except one cluster within the correct diagnostic group. This is the same sample that...
was incorrectly classified in the PLS model (Fig. 2a). In addition, one can observe that proteins of same identity tend to cluster together, e.g. the two gelsolin fragments and the five hemopexin spots representing the full-length protein.

4. Discussion

Protein analysis employing classical proteomics combined with multivariate modeling and identification using mass spectrometry resulted in the discovery of three differentially expressed proteins or their fragments in PPS samples as compared to in controls. This firstly suggests that these proteins may exert key roles in PPS pathophysiology. Secondly, these proteins and their fragments represent potential candidates as biomarkers for the disease. To merit as true biomarkers studies will be required in larger materials of PPS and a variety of other CNS diseases. Notably, however, in comparison with samples from SPMS (being an age-matched control group with ongoing inflammation and neuronal destruction), the most predictive proteins were specific for PPS. In view of the protein expression observed herein in PPS and the functional role of these proteins and potential reasons for their fragmentations, as discussed below, the data strongly support the occurrence of ongoing exaggerated nervous tissue damage in PPS, the underlying cause of which still remains unclear. However, previous observations of the increased expression of pro-inflammatory cytokines in the CSF and their down-modulation upon IVIG treatment combined with a beneficial effect on clinical measures of the disease [24,38,39], strongly argues for a chronic inflammation causing the nervous tissue damage. The underlying cause for such a chronic insidious inflammation decades after the acute paralytic disease still remains elusive. Although there is no unequivocal demonstration of latent poliovirus in PPS, this possibility cannot be excluded. Another possible explanation could be an aberrant, perhaps autoimmune response to nervous tissue, once and originally triggered by the infection.

The increase of kallikrein 6 levels can be regarded as being part of a CNS innate immunity response to a variety of insults, both traumatic and primary inflammatory events [40]. Kallikrein 6 is normally expressed in neurons and oligodendrocytes, and is strongly up-regulated in astrocytes and macrophages/microglia during CNS pathologies such as spinal cord trauma and in human or experimental inflammatory demyelinating diseases [40–42]. It has also been detected in the CSF and CNS of neurendegenerative diseases such as Alzheimer’s and Parkinson’s disease [43]. Our finding of an increase in kallikrein 6 in the CSF of persons with PPS provides a strong additional argument for an ongoing nervous tissue damaging process underlying the syndrome. In view of its induction in lesions of primary CNS inflammatory diseases [42,44], we speculate that pro-inflammatory cytokines such as TNF-α and IFN-γ may also be considered to be involved, in particular in view of their known increased production in PPS [24,38,45]. As with most inflammatory mediator and effector molecules, both beneficial and detrimental effects can be postulated for kallikrein 6. On the potential beneficial side, the enzyme modifies the conditions for neurite outgrowth after spinal cord trauma [40]. On the detrimental side, the enzyme mediates toxicity against myelin/oligodendrocytes in experimental autoimmune encephalomyelitis [42]. Thus while the exact consequences of the observed increased Kallikrein 6 levels in PPS are as yet unclear, it may play a central role in the late symptomatology subsequent to polio.

In the present study fragments of gelsolin were the most predictive among the proteins present in the CSF of persons with PPS. It has been reported that caspase 3 cleaves gelsolin in the pro-apoptotic pathway [46]. Furthermore, TNF-α induces caspase 3-mediated gelsolin cleavage in muscle cells [47]. In its anti-apoptotic role, gelsolin can prevent apoptosis via the mitochondrial pathway [48]. Our data suggest that gelsolin is cleaved by caspase-3 in PPS for the following reasons: (i) the cleavage site of caspase 3 produces an acidic protein variant which matches well with observed Mr/pI, and (ii) all four identified tryptic peptides are located within the same acidic protein (Table 2). Hence the detected fragmentation of gelsolin suggests increased caspase 3 activity during PPS and consequently an increased intrathecal apoptotic activity.

We observed a decrease in levels of the full-length hemopexin protein, along with increased levels of several hemopexin-fragments in CSF. Due to the well-known fact that hemopexin migrates as multiple glycosylated variants in a 2D-gel separation [49], we conclude that the observed parallel decrease of five full-length hemopexin variants represents highly reliable data. The most well characterized function of hemopexin is to bind heme and to prevent against oxidative stress through receptor-mediated induction of heme oxygenase [50]. Interestingly, the role of hemopexin seems to be different in CNS as compared to its systemic functions. It has been reported that hemopexin can be induced in response to tissue damage in the CNS [51]. Changes in hemopexin expression patterns have been recorded in an inflammatory kidney disease, Minimal Change Disease (MCD), and plasma samples from such persons exhibited proteolytic activity resulting in degradation of extracellular matrix [52]. Our data are consistent with an increased proteolytic activity affecting hemopexin in the CNS of persons with PPS, perhaps being a secondary event to chronic inflammation. However, the detailed reason for the fragmentation of hemopexin in PPS is unknown and remains to be further investigated. It should be noted that full-length gelsolin and hemopexin have previously been proposed as biomarker candidates in neurodegenerative diseases [53,54] and that our observations focus on the fragmentation of these proteins. To further evaluate hemopexin and gelsolin as biomarkers we therefore suggest western blot analysis rather than ELISA.

CSF mononuclear mRNA expression of TNF-α and IFN-γ, being two key pro-inflammatory cytokines, have been assessed in all individuals studied presently in a previous study and all had elevated levels [24,38,45]. An increase of CSF TNF-α protein was recently confirmed in the CSF of
PPS patients [18]. This suggests that pro-inflammatory cytokines may be key players in the development and maintenance of PPS. The role of an inflammatory process is further supported by previous findings of inflammatory infiltrates in the spinal cord of post-mortem PPS material [55–57], as well as through signs of inflammation in affected muscles [7,58].

Since we have not had access to CSF from persons with prior poliomyelitis, but with no signs of increasing late symptomatology as in PPS, we cannot know if the observed aberrations are characteristic for prior poliomyelitis overall, or characteristic of the late deterioration. Studies of persons with stable post-polio disabilities are therefore warranted. However, we hypothesize that the inflammatory changes and protein aberrations could be pathogenically relevant even if they would be present in prior poliomyelitis without apparent late sequelae. Different individuals could well be differentially sensitive for chronic low-grade CNS inflammation until critical thresholds are reached, resulting in the varieties of late sequelae. Furthermore, there is experimental evidence in the rat species for genetically determined differences in the susceptibility of motor neurons to non-specific and inflammatory damage, as rat non-MHC genes drastically affect nerve avulsion-induced motor neuron degeneration and CNS innate immunity [59–61]. If there are similar genes in action in the human species, the outcome in the form of neuronal damage may well vary between individuals at similar levels of chronic inflammation as are apparent late after paralytic poliomyelitis.

In conclusion, we herein demonstrate a protein profile, based on its high predictive value, has the potential to serve as a diagnostic biomarker for PPS. The proteins identified in this study are known to be involved in different pathways associated with tissue damage and apoptosis. These data and previous observations of inflammation and cytokine production provide strong support for the hypothesis that PPS is caused by an active inflammatory and neurodegenerative process. There is consequently potential for various modes of anti-inflammatory and/or neuroprotective therapy.

Role of the funding source

The sponsor, AstraZeneca R&D, Sweden, did not interfere with the study design or interpretation of the data nor in the decision to submit the paper for publication. The corresponding author, as well as all other authors, have full access to all the data in the study and have final responsibility for the decision to submit for publication.

Conflict of interest

There are no other known conflicts of interest besides a patent application (reference number US 60/944,816) entitled “Diagnostic Method”. This application was filed 2007-06-19 in US.

Acknowledgements

The proteomics have been done at AstraZeneca R&D, the clinical parts, and sample collections have been supported by grants from the Gustaf V foundation and the Swedish medical research council.

The Swedish research council, Gustav V foundation.
Hugh Salter and Jenny Flygare for reviewing the manuscript. Jon D Randall for technical assistance and support regarding the PDQuest software. Lisbet Broman for her excellent coordination of the clinical data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jprot.2008.11.014.

REFERENCES


Please cite this article as: Gonzalez H, et al. Identification of novel candidate protein biomarkers for the post-polio syndrome — Implications for diagnosis, neurodegeneration and neuroinflammation, J Prot (2009), doi:10.1016/j.jprot.2008.11.014


