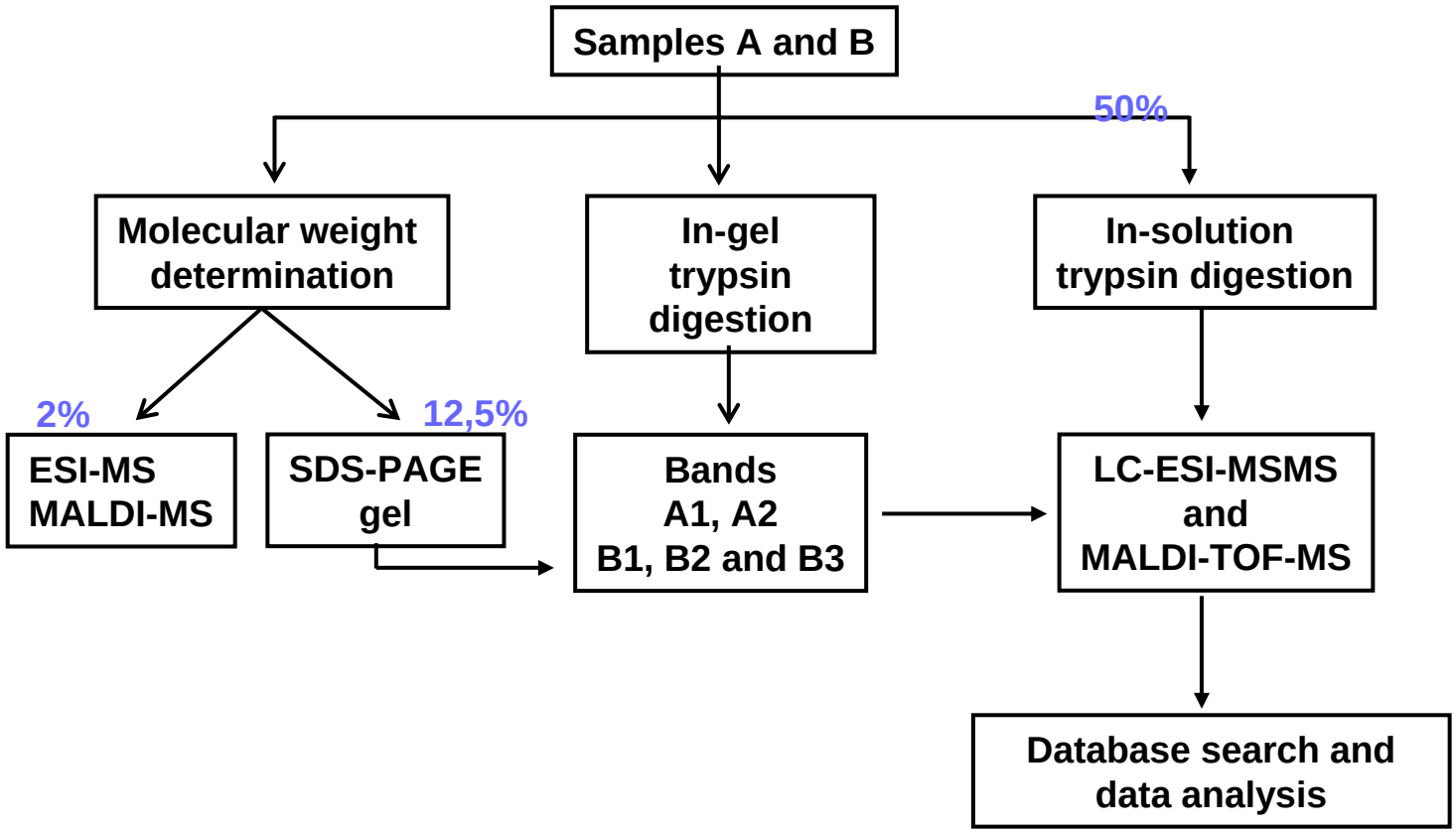
47886

Characterization of ABRF 2008 samples A and B



Protein characterization

We firstly performed an in-solution digestion of samples A and B in order to identified the proteins in the samples, and we also determined the MW of intact proteins via MS. Once we had identified the protein Q15109 in both samples and we saw differences in the MS from samples A and B, we decided to run a SDS-PAGE gel in order to better characterize the bands separated in the gel (A1 and A2 for the sample A and B1, B2 and B3, for the sample B).

RESULTS

- MW Determination:

Sample A: 33,158Da and 33,190Da, respectively by ESI and MALDI-MS.
40KDa and 32KDa, approximately by SDS-PAGE

Sample B: 33,158Da; 24,577Da and 33,178Da; 24,581Da; respectively by ESI and MALDI-MS.
40KDa, 32KDa and 24KDa, approximately by SDS-PAGE

- In-solution and in-gel trypsin digestion results:

1 protein was found in **sample A**: Q15109|RAGE_HUMAN (Advanced glycosylation end product-specific receptor precursor)
UniProt-SwissProt noted sequence:

```
  1 MAAGTAVGAW VLVLSLWGAV VGAQNITARI GEPLVLKCKG APKKPPQRLE
 51 WKLNTGRTEA WKVLSPQGGG PWDSVARVLP NGSLFLPAVG IQDEGIFRCQ
101 AMNRNGKETK SNYRVRVYQI PGKPEIVDSA SELTAGVPNK VGTCVSEGSY
151 PAGTLSWHLD GKPLVPNEKG VSVKEQTRRH PETGLFTLQS ELMVTPARGG
201 DPRPTFSCSF SPGLPRHRAL RTAPIQPRVW EPVPLEEVQL VVEPEGGAVA
251 PGGTVTLTCE VPAQPSPQIH WMKDGVPLPL PPSPVLILPE IGPQDQGTYS
301 CVATHSSHGP QESRAVSISI IEPGEEGPTA GSVGGSGLGT LALALGILGG
351 LGTAALLIGV ILWQRRQRRG EERKAPENQE EEEERAELNQ SEEPEAGESS
401 TGGP
```

Red: peptides found by MSMS

Green: N-terminal peptide + GSHM

Blue: peptides found only by PMF

We found the following differences at the N- and C-terminal in comparison to the noted sequence:
The **N-terminal** is: **GSHMAQNITAR**, where GSHM is the sequence retained with the protein sequence after trombin cleavage site in the His-tag. This peptide sequence was found and manually confirmed using MASCOT, Peaks Studio v4.5 and the “Spider” tool in Peaks Studio v4.5. Moreover, the signal corresponding to the molecular weight of this peptide (1185.58 Da) was found in the mass fingerprinting of both, in-solution and in-gel digestion of sample A.
The C-terminal is truncated in comparison to the noted protein sequence. The C-terminal peptide of the noted protein is TGGP, whereas the homologue protein found in sample A has the C-terminal peptide **AVSISIIEPGEEG** instead. Since the protein was digested with trypsin this data confirm that this is the C-terminal peptide.
MW observed found in both MS analysis of the intact protein and the SDS-PAGE gel would also be consistent with a protein with such C- and N-terminals
We could not explain the difference in MW of bands A1 and A2 observed in the SDS-PAGE.

2 protein were found in **sample B**: Q15109|RAGE_HUMAN (Advanced glycosylation end product-specific receptor precursor) UniProt-SwissProt noted sequence:

```
  1 MAAGTAVGAW VLVLSLWGAV VGAQNITARI GEPLVLKCKG APKKPPQRLE
 51 WKLNTGRTEA WKVLSPQGGG PWDSVARVLP NGSLFLPAVG IQDEGIFRCQ
101 AMNRNGKETK SNYRVRVYQI PGKPEIVDSA SELTAGVPNK VGTCVSEGSY
151 PAGTLSWHLD GKPLVPNEKG VSVKEQTRRH PETGLFTLQS ELMVTPARGG
201 DPRPTFSCSF SPGLPRHRAL RTAPIQPRVW EPVPLEEVQL VVEPEGGAVA
251 PGGTVTLTCE VPAQPSPQIH WMKDGVPLPL PPSPVLILPE IGPQDQGTYS
301 CVATHSSHGP QESRAVSISI IEPGEEGPTA GSVGGSGLGT LALALGILGG
351 LGTAALLIGV ILWQRRQRRG EERKAPENQE EEEERAELNQ SEEPEAGESS
401 TGGP
```

Red: peptides found by MSMS

Green: N-terminal peptide + GSHM

Blue: peptides found only by PMF

Pink: C-terminal of protein B3, found by MSMS

We found two proteins with the same entry number, but with two different C-terminals. In both in-solution and in-gel trypsin digestions we observed the same N-terminal peptide, as described for sample A. As explained above, the N-terminal is GSHMAQNITAR, where GSHM is the sequence retained with the protein sequence after trombin cleavage site in the His-tag. Such sequence was determined using the “Spider” tool in Peaks Studio v4.5.

The two C-terminals found are truncated in comparison to the noted protein sequence. The C-terminal peptide of the noted protein is TGGP, whereas one of the homologues protein found in sample B has the C-terminal peptide AVSISIIEPGEEG instead (bands B1 and B2 in the SDS-PAGE gel). The second C-terminal peptide is VWEPVPLEE (B3 in the gel). Since the protein was digested with trypsin these data confirm that these are the C-terminal peptides.

We could not determine sequence differences between bands B1 and B2, although we saw differences in MW.

The mass of the proteins described do not agree with the masses found by intact mass determination. The theoretical mass of the sequence from GSHM to GEEG would be 33489 Da instead of the 33158 Da determined by ESI-MS, which leads us to think that there are other differences between proteins in the sample and those noted in the database, such as deletion of a couple or more residues, that we were not able to determine, and could explain the 331 Da mass difference.