

Date: Juny 16 , 2006	PROT@GEN [®]	Sample: Endothelial cell line, treated
Project: Cardiovascular disease	Protein Identification Summary Spot: 124	Gel: HUVEC No. 66

Protein 1

Accession: gi|21040384
Proteinname: Vimentin [Homo sapiens]
Taxonomy: Homo sapiens

Protein sequence (466 aa)

```

10          20          30          40          50
MSTRSVSSSS YRRMFGGPGT ASRPSSRSY VTTSTRRTYSL GSALRPSTR
60          70          80          90         100
SLYASSPGGV YATRSSAVRL RSSVPGVRLL QDSVDFSLAD AINTEFKNTR
110         120         130         140         150
TNEKVELQEL NDRFANYIDK VRFLEQQNKI LLAELEQLKG QGKSRLGDLY
160         170         180         190         200
EEMRELRRQ VDQLTNDKAR VEVERDNLAE DIMRLREKLQ EEMLQREEAE
210         220         230         240         250
NTLQSFRQDV DNASLARLDL ERKVESLQEE IAFLKKLHEE EIQELQAQIQ
260         270         280         290         300
EQHVQIDVDV SKPDLTAALR DVRQOYESVA AKNLQEAEEW YKSKFADLSE
310         320         330         340         350
AANRNNDALR QAKQESTEYR RQVQSLTCEV DALKGTNESL ERQMREMEEN
360         370         380         390         400
FAVEAANYQD TIGRLQDEIQ NMKEEMARHL REYQDLLNVK MALDIEIATY
410         420         430         440         450
RKLLEGEESR ISLPLPNFSS LNLRETNLDS LPLVDTHSKR TLLIKTTVETR
460
DGQVINETSQ HHDDLE

```

Protein properties

MW - seq.	53.7 kDa	pI - seq.	5.1
MW - gel	75.0 kDa	pI - gel	5.0

Sequence coverage

Total	MS (PMF)	MS/MS (PFF)
244 / 466 aa (52.4 %)	244 / 466 aa (52.4 %)	90 / 466 aa (19.3 %)

Date: Juny 16 , 2006	PROT@GEN [®]	Sample: Endothelial cell line, treated
Project: Cardiovascular disease	Protein Identification Summary Spot: 124	Gel: HUVEC No. 66

Alias proteins

Proteins with identical sequence

Accession	Source db	Reference	Name
gi 21040384	gb	AAH30573.1	Vimentin [Homo sapiens]
gi 13111801	gb	AAH00163.2	Vimentin [Homo sapiens]
gi 13629179			vimentin [Homo sapiens]
gi 14742600			(XM_042950) vimentin Homo sapiens
gi 14742608			vimentin [Homo sapiens]
gi 20473667	ref	XP_167414	similar to Vimentin Homo sapiens
gi 2119204	pir		A25074 vimentin - human
gi 22854562	gb	AAN09720.1	CTCL tumor antigen HD-CL-06 [Homo sapiens]
gi 37850	emb	CAA39600.1	unnamed protein product [Homo sapiens]
gi 44890587	gb	AAH66956.1	Vimentin [Homo sapiens]
gi 55977767	sp	P08670	VIME_HUMAN Vimentin
gi 62414289	ref	NP_003371.2	vimentin [Homo sapiens]
gi 62896573	dbj	BAD96227.1	vimentin variant [Homo sapiens]
gi 62896763	dbj	BAD96322.1	vimentin variant [Homo sapiens]
gi 7576229	emb	CAB87963.1	vimentin [Homo sapiens]

Spots with same identified protein in this project

Sample	Gel	Spot
Endothelial cell line, untreated	HUVEC No. 67	132

MS (PMF)

ProFound Score	Mascot Score	MS-Fit Score	MetaScore	Int. Coverage	Int. Cov. Calibrants	Database	Spectrum
2.4	238	-	99.2	47.5 %	15.9 %	NCBIInr	1/1SRef

MS/MS (PFF)

Parent Mass	Sequest MetaScore	Mascot Score	PFF-SolverExp	Sequence	Database	Spectrum
-	53.7	-	-	EMEENFAVEAANYQDTIGR FADLSEANR FADLSEANRNNDALR ISLPLPNFSSLNLR LGDLYEEEMR LLEGESR NLQEAEEWYK TNEKVELQELNDR VELQELNDR	NCBIInr	Combined PFF

What is displayed in the Protein Identification Summary?

Protein 1

In case of several identified proteins the proteins are numbered consecutively.

Accession

Accession number of the identified protein in the used sequence database.

Proteinname

Name of the identified protein as given in the sequence database.

Taxonomy

Organism of the identified protein as given in the sequence database (sometimes not given by the sequence database).

Protein sequence

Amino acid sequence of the identified protein and total number of amino acids. The meaning of the highlighted areas within the amino acid sequence is as follows:

The light red highlighted sequence areas indicate peptides which are assigned by peptide mass fingerprint (PMF resp. MS). The dark red highlighted sequence areas indicate peptides which are assigned by peptide fragmentation fingerprint (PFF resp. MS/MS).

Protein properties

Comparison of molecular weight (MW) [kDa] and pI:

- a) calculated from the amino acid sequence (**MW – seq. and pI – seq.**)
- b) estimated from the gel position (**MW – gel and pI – gel**)

Sequence coverage

The sequence coverage is defined as the ratio of the number of the explained amino acids to the total number of amino acids of the protein [%]. It is given in **total** as well as for **PMF**- and **PFF**-analysis separately.

Alias proteins

One or more proteins related to the identified protein are listed in two different groups:

Proteins with identical sequence

Entries of the used sequence database which have the identical amino acid sequence (100%) but different accession numbers (**Accession**) (→ redundant entries). The table

also shows the sequence database (**Source DB**) where the entry originally comes from and the corresponding reference number (**Reference**) and protein name (**Name**).

Spots with same identified protein in this project

Spots within the project which were identified as the same protein. The corresponding **Sample**, **Gel**, and spot name (**Spot**) is given.

MS (PMF)

Parameters for the evaluation of the result(s) of the peptide mass fingerprint (PMF):

ProFound Score

Describes the significance of the search result from the search engine ProFound. The threshold value is 1.65. All results with values ≥ 1.65 have a 95%age probability not to be false positive, for values ≥ 2.3 this probability is 99%.

Lit: Zhang W, Chait BT. *Anal Chem.*, 2000 Jun 1; 72(11): 2482-9.

Mascot Score

Describes the significance of the search result from the search engine Mascot. There is no absolute threshold value for Mascot. It depends on the search parameters, especially on the size of the used sequence database. It is calculated and given for each search. A significance threshold of about 60 is typical. This is equivalent to a 95%age probability that the result is not false positive.

Lit: Perkins DN, Pappin DJ, Creasy DM, Cotrell JS. *Electrophoresis*, 1999; 20(18): 3551-3567.

MS-Fit Score

Describes the significance of the search result from the search engine MS-Fit. The exponential progression of the value makes a correlation of a value to a false positive rate difficult. There is no known significant level of MS-Fit.

Lit: Pappin DJC, Hojrup P, Bleasby AJ. *Curr. Biol.*, 1993; 3(6): 327-32.

MetaScore

By using more than one search engine the significance of the search result is increased. To enable the evaluation of this significance a score was generated which combines several search results. The 5% significance level of the *MetaScore* was empirically determined to be 80. That means, the probability to obtain a complete random result with a value ≥ 80 is less than 5%.

Lit: Chamrad DC, Koerting G, Gobom J, Thiele H, Klose J, Meyer HE, Blueggel M. *Anal Bioanal Chem.*, 2003 Aug; 376(7): 1014-22.

Int. Coverage

The intensity coverage is defined as the ratio of the accumulated intensities of the peaks explained by peptides of the identified protein to the total intensity of all peaks of the PMF spectrum [%].

Int. Cov. Calibrants

Every PMF spectrum contains known peptide masses, for example trypsin autolyse peaks. These known masses are used for calibrating the spectrum and are therefore named as calibrants. The intensity coverage of the calibrants is defined as the ratio of the accumulated intensities of the calibrant peaks to the total intensity of all peaks of the PMF spectrum [%].

Database

The sequence database which was used for the search.

Spectrum

Name of the PMF spectrum.

MS/MS (PFF)

Parameters for the evaluation of the result(s) of peptide fragmentation fingerprints (PFF):

Parent Mass

Parent mass of the fragmented peptide. If several PFF spectra were combined for the search (see below) the parent masses are not shown.

Sequest MetaScore

If several PFF spectra are acquired for one protein spot routinely a database search is performed which combines all acquired PFF spectra. Thus the significance of the search result is increased. To enable the evaluation of this significance the *Sequest MetaScore* was generated which combines the SEQUEST search results of all PFF spectra.

Mascot Score

Describes the significance of the search result from the search engine Mascot. Both the values for the single and for the combined PFF searches are shown in this column.

Lit: Perkins DN, Pappin DJ, Creasy DM, Cotrell JS. *Electrophoresis*, 1999; 20(18): 3551-3567.

PFF-Solver Exp.

Describes the significance of the search result from the search engine PFF-Solver. The 5% significance level is set to Exp. 95. Therefore the expectation value for single PFF searches corresponds to the probability roughly. That means, the probability to obtain a complete random result with an Exp. value ≥ 95 is less than 5%. For Exp. values ≥ 99

this probability is less than 1%. Both the values for the single and for the combined PFF searches are shown in this column.

Sequence

Amino acid sequences of the fragmented and identified peptides.

Database

The sequence database which was used for the search.

Spectrum

Name of the PFF spectrum.