Characterisation of antibodies to Migration Stimulating Factor (MSF). Detection of MSF isoforms.

Schor, Ana; Kankova , Katerina; Woolston, Anne-Marie; Vojtesek, Borivoj; Felts, Paul; Norman, David

DOI:
10.20933/100001167

Publication date:
2020

Document Version
Other version

DRAFT. REPORT TO BE COMPLETED

Title: Characterisation of antibodies to Migration Stimulating Factor (MSF). Detection of MSF isoforms.

Authors: Schor AM[(1)], Kankova K[(2)], Woolston AM[(3)], Vojtesek B[(4)], Felts PA[(1)], Norman DG[(5)], Harada K[(6)]

(1) School of Science and Engineering, University of Dundee, Dundee, DD1 4HN, UK

(2) Department of Pathophysiology, Faculty of Medicine, Masaryk University, Kamenice 5, Brno, Czech Republic

(3) School of Dentistry, University of Dundee, Dundee, DD1 4HR, UK

(4) RECAMO, Masaryk Memorial Cancer Institute, Zluty kopec 7, 656 53 Brno, Czech Republic

(5) School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK

(6) Department of Oral and Maxillofacial Surgery, Yamaguchi University Graduate School of Medicine, 1-1-1 Minamikogushi, Ube, Yamaguchi 755-8505, Japan

https://doi.org/10.20933/100001167

Time period during which data was collected: 1994-2019

Abstract

Migration Stimulating Factor (MSF) is a 70kDa truncated isoform of fibronectin (FN). Unlike FN, MSF is not a matrix molecule but a soluble factor which exhibits a range of potent cytokine-like bioactivities not displayed by full-length FN. Two isoforms of human MSF (MSF+aa and MSF-aa), as well as murine Migration Stimulating Factor (mMSF) have been cloned. In this communication we report the characterisation of various polyclonal and monoclonal antibodies to human and murine MSF. In particular:

(i) Specific MSF-identification antibodies that recognise both MSF+aa and MSF-aa;

(ii) Specific mMSF-identification antibodies;

(iii) Identification antibodies that recognise MSF-aa but not MSF+aa;

(iv) MSF-function-neutralising antibodies that recognise MSF+aa, MSF-aa, mMSF and the gelatin-binding domain of FN/MSF (Gel-BD) but not full-length FN.

Description of the antigens:

Fibronectin (FN) is a modular glycoprotein consisting of the following functional domains: Hep 1/Fib-1 (N-terminal low affinity binding to heparin and fibrin), Gel-BD (binding to gelatin/collagen), Cell-BD (RGD-mediated binding to integrins), Hep-2 (high affinity heparin binding) and Fib-2 (C-terminal fibrin binding site). Each functional domain is composed of a different number of three possible homology modules, called type I, II and III (**Fig 1**). Human Migration Stimulating Factor (referred to as MSF) is a 70kDa truncated isoform of FN. MSF RNA is generated from the fibronectin gene by a two-stage processing mechanism. In the first stage, an MSF-specific primary transcript is generated from the fibronectin gene by read-through of intron 12, separating exons III-1a and III-1b. This is followed by intra-intronic cleavage to produce a 5.9 kb MSF pre-message that remains sequestered within the nucleus, where it is rapidly degraded. In cells that express MSF protein a second stage takes place, whereby the intron-derived 3' UTR of the pre-message is cleaved a second time to produce a 2.1 kb mature MSF message. This has a shorter (195bp) intron-derived 3' sequence containing a 30bp in-frame coding sequence (immediately contiguous with exon III-1a), followed by a 165bp 3'-UTR containing several in-frame stop codons and a cleavage/polyadenylation signal. The mature message is rapidly exported to the cytoplasm for translation [Schor et al, 2003; Kay et al, 2005]. Therefore, MSF is identical to the N terminus of full-length fibronectin, up to and including the amino acid sequence coded by exon III-1a, with the addition of an MSF-unique (intron-coded) 10 amino acid C-terminus: VSIPPRNLGY [Schor et al, 2003; Kay et al, 2005] (**Fig 1**).

FN is a major component of the extracellular matrix. Unlike FN, MSF is a monomer and is not a matrix molecule but a soluble factor which exhibits a range of cytokine-like bioactivities, including the stimulation of cell migration and angiogenesis. The motogenic activity of MSF is mediated by the IGD motifs, present in modules I3, I5, I7 and I9 of the Gel-BD domain and, in some cases, by the HEEGH motif present in module I8 [Schor et al, 1999, 2003; Houard et al, 2005; Millard et al, 2007] (**Fig 1, Fig 2**). The bioactivities of MSF are not expressed by full-length fibronectin, due to steric hindrance [Millard et al, 2007; Vakonakis et al 2009].

Two isoforms of MSF (referred to as MSF) have been cloned. Both contain the same unique 10 amino acid C-terminus as well as same bioactive IGD and HEEGH amino acid motifs. The two isoforms differ solely in terms of a 45bp deletion in exon II-1 and are consequently referred to as MSF+aa and MSF-aa to indicate the retention or deletion of a 15 amino acid

sequence in module II-1 (**Fig 1** and **Fig 2**). The term MSF or total MSF will be employed to denote both isoforms.

We have isolated and cloned a murine MSF (mMSF) transcript by PCR. This is homologous to its human counterpart, consisting of the 5-terminus of mouse FN, up to and including exon III-1a, and terminating in a unique 3'coding sequence derived from the intron separating exons III-1a and -1b. The 3'UTR ends in a polyA tail. mMSF protein consequently has a molecular mass of 70kDa and terminates in a unique 12 amino acid C-terminus: VSNSSAALDSDP (**Fig 3**). The murine FN coding sequence is over 90% homologous to human FN; the four IGD motifs and the HEEGH motif are similarly located in modules I3, I5, I7, I9 and I8, respectively. However, there is no significant homology between the human and mouse intron-derived C-terminal MSF-unique peptides.
Eukaryotic and prokaryotic recombinant MSF and mMSF were produced as described [Schor et al 2003].
We have raised polyclonal (Pab) and monoclonal (Mab) antibodies to human and murine MSF. The peptides used as antigens to raise these antibodies and an overview of the results obtained are shown in **Table 1**. The antibodies were characterised by ELISA, immunoblotting, IHC and their ability to abrogate or remove MSF/mMSF bioactivity [Schor et al 2003, 2012]. As expected, different antibodies were useful for certain techniques and not for others.

Production of antibodies

To be completed

Characterisation of VSI antibodies

To be completed

Conclusions: VSI are MSF-specific identification antibodies that recognise the unique 10-mer C-terminal sequence of MSF+aa and MSF-aa; that is, total MSF.

Characterisation of TYN antibodies

To be completed

Conclusions: TYN are identification antibodies that recognise MSF-aa but not MSF+aa.

Characterisation of VSN antibodies

To be completed

Conclusions: VSN are mMSF-specific identification antibodies that recognise the unique 12-mer C-terminal sequence of mMSF.

Characterisation of pepQ antibodies

To be completed

Conclusions: pepQ are MSF-function-neutralising antibodies that recognise MSF+aa, MSF-aa, mMSF and Gel-BD but not FN.

Methods

To be completed

1. ELISA
2. Dot Blots
3. Western Blots
3. Immunohistochemistry (IHC)
4. Cell migration
5. Cell proliferation

ACKNOWLEDGEMENTS

This work has been funded by Cancer Research UK, Breast Cancer Campaign, Engineering and Physical Sciences Research Council, Biotechnology and Biological Sciences Research Council, Scottish Chief Scientist Office and Medical Research Council. We thank the many collaborators that have contributed to this work by providing specimens and experimental data.

Table 1. Overview of the antibodies raised and results obtained.

The peptides indicated were used as antigens to raise monoclonal (Mab) and polyclonal (Pab) antibodies.

Peptide used as antigen	Ab code	Reactivity of Abs
VSIPPRNLGY 10 mer, MSF-unique C-terminus	VSI	Mab and Pab recognise MSF+aa and MSF-aa. Do not recognise FN, Gel-BD or Hep 1/Fib-1 domains.
TYNDRTDSTTSNY 13 mer, present in MSF-aa, II-1. In MSF+aa these amino acids are adjacent to the sequence deleted in MSF-aa (6 before and 7 after)	TYN	Mab and Pab recognise MSF-aa. Do not recognise MSF+aa, FN or Gel-BD
VSNSSAALDSDP 12 mer, mMSF-unique C-terminus	VSN	Pab recognise mMSF. Do not recognise MSF, FN or Gel-BD
TNEGVMYRIGDQWDKQHDMGH 21-mer, IGD-containing peptide in module I-7	pepQ	Mab recognise MSF+aa, MSF-aa and Gel-BD. Do not recognise FN.

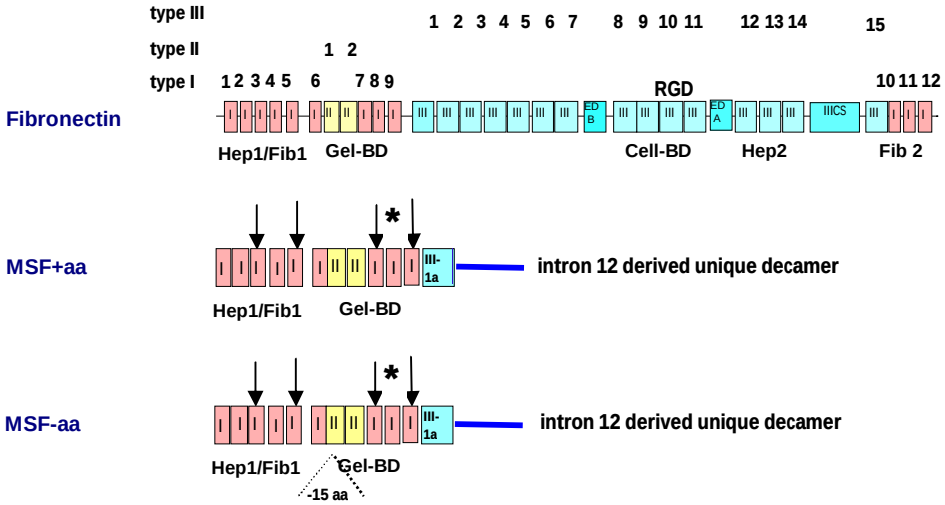

Figure 1. The structure of fibronectin and MSF. Fibronectin domains include: Hep 1/Fib-1 (N-terminal low affinity binding to heparin and fibrin), Gel-BD (binding to gelatin/collagen), Cell-BD (RGD-mediated binding to integrins), Hep-2 (high affinity heparin binding) and Fib-2 (C-terminal fibrin binding site). Each domain is composed of three possible homology modules, called type I, II and III. MSF is identical to the N terminus of fibronectin, up to and including the amino acid sequence coded by exon III-1a, with the addition of an MSF-unique (intron-coded) 10 amino acid C-terminus. Two isoforms of MSF have been cloned. These differ solely in terms of a 45bp deletion in exon II-1 and are consequently referred to as MSF+aa and MSF-aa to indicate the retention or deletion of a 15 amino acid sequence in module II-1. The location of IGD motifs (↓) and HEEGH motif (*) is indicated.

MSF+aa. Accession number AJ535086
http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=Protein&id=27227743

```
1   MLRGPGPGLL LLAVQCLGTA VPSTGASKSK RQAQQMVQPQ SPVAVSQSKP GCYDNGKHYQ
61  INQQWERTYL GNALVCTCYG GSRGFNCESK PEAEETCFDK YTGNTYRVGD TYERPKDSMI
121 WDCTCIGAGR GRISCTIANR CHEGGQSYKI GDTWRRPHET GGYMLECVCL GNGKGEWTCK
181 PIAEKCFDHA AGTSYVVGET WEKPYQGWMM VDCTCLGEGS GRITCTSRNR CNDQDTRTSY
241 RIGDTWRKKD NRGNLLQCIC TGNGRGEWKC ERHTSVQTTS SGSGPFTDVR AAVYQPQPHP
301 QPPPYGHCVT DSGVVYSVGM QWLKTQGNKQ MLCTCLGNGV SCQETAVTQT YGGNSNGEPC
361 VLPFTYNGRT FYSCTTEGRQ DGHLWCSTTS NYEQDQKYSF CTDHTVLVQT RGGNSNGALC
421 HFPFLYNNHN YTDCTSEGRR DNMKWCGTTQ NYDADQKFGF CPMAAHEEIC TTNEGVMYRI
481 GDQWDKQHDM GHMMRCTCVG NGRGEWTCIA YSQLRDQCIV DDITYNVNDT FHKRHEEGHM
541 LNCTCFGQGR GRWKCDPVDQ CQDSETGTFY QIGDSWEKYV HGVRYQCYCY GRGIGEWHCQ
601 PLQTYPSSSG PVEVFITETP SQPNSHPIQW NAPQPSHISK YILRWRPVSI PPRNLGY
```

MSF-aa: Accession number AJ276395 (15 amino acid deletion in module II-1)
http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=Protein&id=12053817

```
1   MLRGPGPGLL LLAVQCLGTA VPSTGASKSK RQAQQMVQPQ SPVAVSQSKP GCYDNGKHYQ
61  INQQWERTYL GNALVCTCYG GSRGFNCESK PEAEETCFDK YTGNTYRVGD TYERPKDSMI
121 WDCTCIGAGR GRISCTIANR CHEGGQSYKI GDTWRRPHET GGYMLECVCL GNGKGEWTCK
181 PIAEKCFDHA AGTSYVVGET WEKPYQGWMM VDCTCLGEGS GRITCTSRNR CNDQDTRTSY
241 RIGDTWSKKD NRGNLLQCIC TGNGRGEWKC ERHTSVQTTS SGSGPFTDVR AAVYQPQPHP
301 QPPPYGHCVT DSGVVYSVGM QWLKTQGNKQ MLCTCLGNGV SCQETAVTQT YGGNSNGEPC
361 VLPFTYNDRT DSTTSNYEQD QKYSFCTDHT VLVQTRGGNS NGALCHFPFL YNNHNYTDCT
421 SEGRRDNMKW CGTTQNYDAD QKFGFCPMAA HEEICTTNEG VMYRIGDQWD KQHDMGHMMR
481 CTCVGNGRGE WTCIAYSQLR DQCIVDDITY NVNDTFHKRH EEGHMLNCTC FGQGRGRWKC
541 DPVDQCQDSE TGTFYQIGDS WEKYVHGVRY QCYCYGRGIG EWHCQPLQTY PSSSGPVEVF
601 ITETPSQPNS HPIQWNAPQP SHISKYILRW RPVSIPPRNL GY
```

Figure 2. MSF+aa and MSF-aa protein sequences. The sequences highlighted are: MSF-unique decamer in red (sequence MSF does not share with FN). MSF 15 amino acid region present in MSF+aa and absent in MSF-aa in blue. IGD and HEEGH motifs in purple.

```
   1   M  L  R  G  P  G  P  G  R  L  L  L  L  A  V  L  C  L  G  T
   1  ATGCTCAGGGGTCCGGGACCCGGGCGGCTGCTGCTGCTGGCAGTCCTGTGCCTGGGGACC
  21   S  V  R  C  T  E  A  G  K  S  K  R  Q  A  Q  Q  I  V  Q  P
  61  TCGGTGCGCTGCACCGAAGCCGGGAAGAGCAAGAGGCAGGCTCAGCAAATCGTGCAGCCT
  41   Q  S  P  V  A  V  S  Q  S  K  P  G  C  F  D  N  G  K  H  Y
 121  CAATCCCCGGTGGCTGTCAGTCAGAGCAAGCCTGGCTGTTTTGACAATGGGAAGCACTAT
  61   Q  I  N  Q  Q  W  E  R  T  Y  L  G  N  A  L  V  C  T  C  Y
 181  CAGATAAATCAGCAGTGGGAACGGACCTACCTAGGCAACGCCCTGGTTTGTACCTGCTAT
  81   G  G  S  R  G  F  N  C  E  S  K  P  E  P  E  E  T  C  F  D
 241  GGAGGAAGCCGGGGTTTTAACTGCGAGAGCAAGCCTGAGCCTGAAGAGACTTGCTTTGAC
 101   K  Y  T  G  N  T  Y  K  V  G  D  T  Y  E  R  P  K  D  S  M
 301  AAATACACTGGGAACACTTACAAAGTGGGTGACACTTATGAGCGCCCTAAAGATTCCATG
 121   I  W  D  C  T  C  I  G  A  G  R  G  R  I  S  C  T  I  A  N
 361  ATCTGGGACTGTACCTGCATCGGGGCTGGGAGAGGCAGGATCAGCTGTACCATTGCAAAT
 141   R  C  H  E  G  G  Q  S  Y  K  I  G  D  K  W  R  R  P  H  E
 421  CGCTGCCATGAAGGGGGTCAGTCCTACAAGATTGGCGACAAGTGGAGGAGGCCACATGAG
 161   T  G  G  Y  M  L  E  C  L  C  L  G  N  G  K  G  E  W  T  C
 481  ACTGGTGGCTACATGTTAGAGTGTCTGTGTCTGGGAAATGGAAAAGGGGAATGGACCTGC
 181   K  P  I  A  E  K  C  F  D  H  A  A  G  T  S  Y  V  V  G  E
 541  AAACCTATAGCTGAGAAGTGTTTTGATCATGCTGCTGGGACGTCCTACGTCGTGGGGGAG
 201   T  W  E  K  P  Y  Q  G  W  M  M  V  D  C  T  C  L  G  E  G
 601  ACCTGGGAAAAGCCCTACCAAGGCTGGATGATGGTGGACTGTACTTGTCTAGGCGAAGGC
 221   N  G  R  I  T  C  T  S  R  N  R  C  N  D  Q  D  T  R  T  S
 661  AATGGACGCATCACCTGTACCTCCAGAAACAGATGCAACGATCAGGACACCCGGACATCC
 241   Y  R  I  G  D  T  W  S  K  K  D  N  R  G  N  L  L  Q  C  V
 721  TATAGGATTGGAGACACGTGGAGCAAGAAGGACAACCGAGGAAACCTGCTTCAGTGTGTC
 261   C  T  G  N  G  R  G  E  W  K  C  E  R  H  A  L  Q  S  A  S
 781  TGCACAGGCAATGGCAGAGGGGAGTGGAAGTGTGAGCGACATGCTCTACAAAGTGCTTCA
 281   A  G  S  G  S  F  T  D  V  R  T  A  I  Y  Q  P  Q  T  H  P
 841  GCCGGATCTGGCTCCTTCACTGATGTCCGAACAGCTATTTACCAACCGCAGACTCACCCC
 301   Q  P  A  P  Y  G  H  C  V  T  D  S  G  V  V  Y  S  V  G  M
 901  CAGCCCGCTCCCTACGGCCACTGTGTCACCGACAGTGGTGTGGTCTACTCTGTGGGAATG
 321   Q  W  L  K  S  Q  G  N  K  Q  M  L  C  T  C  L  G  N  G  V
 961  CAGTGGCTGAAGTCGCAAGGAAACAAGCAAATGCTGTGCACGTGCCTGGGCAATGGCGTC
 341   S  C  Q  E  T  A  V  T  Q  T  Y  G  G  N  S  N  G  E  P  C
1021  AGCTGCCAGGAGACAGCCGTGACCCAGACTTATGGTGGCAATTCAAACGGGGAGCCCTGT
 361   V  L  P  F  T  Y  N  G  R  T  F  Y  S  C  T  T  E  G  R  Q
1081  GTCCTCCCGTTCACCTACAACGGTAGGACCTTCTATTCCTGCACCACCGAAGGGCGGCAA
 381   D  G  H  L  W  C  S  T  T  S  N  Y  E  Q  D  Q  K  Y  S  F
1141  GACGGACATCTGTGGTGTAGCACAACTTCCAATTACGAACAAGACCAGAAGTATTCCTTC
 401   C  T  D  H  A  V  L  V  Q  T  R  G  G  N  S  N  G  A  L  C
1201  TGCACAGACCATGCGGTTTTGGTTCAGACTCGAGGCGGAAATTCCAATGGTGCTCTGTGC
 421   H  F  P  F  L  Y  N  N  R  N  Y  T  D  C  T  S  E  G  R  R
1261  CACTTCCCCTTCCTGTACAACAACCGGAATTACACCGACTGTACTTCTGAGGGTCGCAGG
 441   D  N  M  K  W  C  G  T  T  Q  N  Y  D  A  D  Q  K  F  G  F
1321  GACAACATGAAATGGTGCGGCACCACCCAGAACTACGATGCCGATCAGAAGTTTGGATTC
 461   C  P  M  A  A  H  E  E  I  C  T  T  N  E  G  V  M  Y  R  I
1381  TGCCCAATGGCTGCCCACGAGGAGATCTGCACAACCAATGAAGGGGTCATGTATCGCATT
 481   G  D  Q  W  D  K  Q  H  D  L  G  H  M  M  R  C  T  C  V  G
1441  GGGGATCAGTGGGATAAGCAGCATGACCTGGGCCACATGATGAGGTGCACGTGTGTGGGG
 501   N  G  R  G  E  W  A  C  I  P  Y  S  Q  L  R  D  Q  C  I  V
1501  AACGGTCGTGGAGAATGGGCCTGCATCCCCTACTCCCAGCTCCGAGACCAGTGCATCGTT
 521   D  D  I  T  Y  N  V  N  D  T  F  H  K  R  H  E  E  G  H  M
1561  GATGACATTACTTACAATGTGAACGACACGTTCCACAAGCGTCACGAGGAGGGACATATG
 541   L  N  C  T  C  F  G  Q  G  R  G  R  W  K  C  D  P  I  D  Q
1621  CTGAACTGTACCTGCTTTGGTCAGGGCCGGGGCAGATGGAAGTGTGACCCCATTGACCAG
 561   C  Q  D  S  E  T  R  T  F  Y  Q  I  G  D  S  W  E  K  F  V
1681  TGCCAAGATTCAGAGACCCGGACATTTTACCAGATTGGTGACTCCTGGGAGAAGTTTGTG
 581   H  G  V  R  Y  Q  C  Y  C  Y  G  R  G  I  G  E  W  H  C  Q
1741  CATGGTGTCCGATACCAGTGTTACTGCTACGGCCGTGGCATCGGGGAGTGGCACTGTCAA
 601   P  L  Q  T  Y  P  G  T  T  G  P  V  Q  V  I  I  T  E  T  P
1801  CCTCTGCAGACCTACCCAGGCACAACTGGACCTGTCCAAGTAATTATCACGGAGACCCCC
 621   S  Q  P  N  S  H  P  I  Q  W  N  A  P  E  P  S  H  I  T  K
1861  AGCCAGCCCAATTCCCACCCCATCCAGTGGAATGCCCCGGAGCCTTCACACATCACCAAG
 641   Y  I  L  R  W  R  P  V  S  N  S  S  A  A  L  D  S  D  P  -
1921  TACATTCTCAGATGGAGACCTGTGAGTAATAGCTCCGCAGCCTTGGACTCTGACCCCTGA
```

Figure 3. Murine MSF protein and nucleotide sequences. The protein sequences highlighted are: MSF-unique decamer in red (sequence mMSF does not share with mouse FN). mMSF 15 amino acid region absent in MSF-aa in blue. IGD and HEEGH motifs in purple.