# Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2 and MDC3: novel human cellular disintegrins highly expressed in the brain

Koji SAGANE<sup>1</sup> Yukio OHYA, Yoshikazu HASEGAWA and Isao TANAKA Tsukuba Research Laboratories, Eisai Co., Ltd., Tokodai 5-1-3, Tsukuba, Ibaraki, Japan

Cellular disintegrins are a family of membrane-anchored proteins structurally related to snake venom disintegrins, and are potential regulators of cell-cell and cell-matrix interactions. The members of this protein family are also called ADAMs (a disintegrin and metalloproteinase) or MDC proteins (metalloproteinase-like disintegrin-like cysteine-rich), because they all contain disintegrinlike and metalloproteinase-like domains. In this paper, we report the cloning and sequence analysis of two novel additional members of this family, which we have termed MDC2 and MDC3. The deduced amino acid sequences reveal that the two proteins possess typical cellular disintegrin-like, cysteine-rich, epidermal growth factor-like, transmembrane, and cytoplasmic

# INTRODUCTION

Cell adhesion molecules are believed to play important roles in neural function and development. Recently, a new class of potential adhesion molecules has been found. The members of this class of proteins are known as cellular disintegrins [also known as ADAMs (a disintegrin and metalloproteinase) or MDCs (metalloproteinase-like, disintegrin-like, cysteine-rich proteins)]. A large number of cellular disintegrins have been detected in various tissues of mammals and lower eukaryotes [1]. The cellular disintegrins have a multiple domain structure: pro-, metalloproteinase-like, disintegrin-like, cysteine-rich, epidermal growth factor-like, transmembrane and cytoplasmic domains. The metalloproteinase-like domain of the cellular disintegrin exhibits sequence similarity with a snake venom haemorrhagic metalloproteinase. It has been demonstrated that some members of this protein family possess metalloproteinase-like domains which are catalytically active and degrade specific substrates [2]. For example, the Drosophila metalloproteinase disintegrin Kuzbanian activates a notch receptor during development [3], and both human MADM/ADAM10 and human tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) converting enzyme (TACE) release soluble TNF- $\alpha$  by proteolysis [4–6]. On the other hand, the disintegrinlike domains of cellular disintegrins exhibit sequence similarity to the snake venom disintegrins and are predicted to function as integrin ligands. It has been revealed that the mouse fertilin  $\beta$ /ADAM2 interacts with the  $\alpha 6/\beta 1$  integrin via its disintegrinlike domain [7].

domains] and exhibit high sequence similarity with human MDC/ADAM11 protein [Katagiri, Harada, Emi and Nakamura (1995) Cytogenet. Cell Genet. **68**, 39–44]. A zinc-binding motif, which is critical for proteinase activity, is disrupted in the metalloproteinase-like domain of MDC2 and MDC3, as well as MDC/ADAM11. In the disintegrin-like domain of snake venom short disintegrins, the RDG-containing loops are critical for integrin binding. These three MDCs do not contain the RDG sequences, but the corresponding loops in these proteins are similar to each other. Northern blot analysis revealed that

Our studies of cellular disintegrins expressed in the brain resulted in the discovery of two novel cDNAs, termed MDC2 and MDC3. Furthermore, we have confirmed that MDC2 and MDC3 mRNA expression was specifically restricted to the brain.

the mRNAs of MDC2, MDC3 and MDC/ADAM11 are highly expressed in the brain. These findings suggest that these proteins

may function as integrin ligands in the brain.

# **EXPERIMENTAL**

# Materials

The Expand High Fidelity PCR system, supplied by Boehringer Mannheim was used for all PCR analyses. SuperScript II reverse transcriptase (Gibco/BRL) was used for all reverse transcription. Human and mouse brain poly(A)<sup>+</sup> RNAs were purchased from Clontech. The 3'-RACE system for rapid amplification of cDNA ends and 5'-RACE system were purchased from Gibco/BRL. The primers, adaptor primer [5'-GGCCACGCGTCGACTAG-TAC(T)17-3'] and universal amplification primer (UAP) (5'-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3'), were included in the 3'-RACE kit. The primers abridged anchor primer (AAP) (5'-GGCCACGCGTCGACTAGTACGGGIIG-GGIIGGGIIG-3') was included in the 5'-RACE kit. A Marathon cDNA amplification kit was purchased from Clontech. The adaptor (Marathon cDNA adaptor) and the primers AP1 and AP2 were included in the kit.

#### **DNA** sequencing

DNA sequencing was performed using a Dye terminator kit and the ABI PRISM 377 DNA sequencer (Perkin–Elmer).

Abbreviations used: MDC protein, metalloproteinase-like disintegrin-like cysteine-rich protein; ADAM, a disintegrin and metalloproteinase; RT, reverse-transcription; RACE, rapid amplification of cDNA ends; EST, expressed sequence tag; TACE, tumour necrosis factor- $\alpha$  converting enzyme; UAP, universal amplification primer; AAP, abridged anchor primer; MTN, multiple tissue Northern.

<sup>&</sup>lt;sup>1</sup> To whom correspondence should be addressed (e-mail k1-sagane@eisai.co.jp).

The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/GenBank Nucleotide Sequence Databases under the accession numbers AB009671 (human MDC2), AB009672 (human MDC3) and AB009675 (human MDC/ADAM11).

#### **Cloning of MDC3 cDNA**

To identify novel cellular disintegrins, we searched the expressed sequence tag (EST) database with the human MDC/ADAM11 [8,9] coding sequences and found that the 340 bp EST sequence HSC2RF011 (Genbank accession number F08148) was highly similar to it. The putative amino acid sequences encoded by the EST sequence suggested that it was the novel disintegrin-like protein. We have henceforth called this sequence MDC3. To clone the full-length cDNA, 3'-RACE was performed using the 3'-RACE system. Human brain  $poly(A)^+$  RNA (0.5 µg) was primed with adaptor primer and reverse transcribed. For the primary PCR reaction, the gene-specific primer MRP001 (5'-TGGAGCCTGCCTTTCCAACAG-3') and the UAP were used and the primers MRP003 (5'-GCTGTAACAATACCTCA-TGTCT-3') and UAP were employed for nested PCR. A 1.3 kb DNA fragment was amplified and subcloned into the pUC18 vector. To get 5'-flanking sequences, 5'-RACE was performed using the Marathon cDNA amplification kit (Clontech). Human brain poly(A)<sup>+</sup> RNA (0.5  $\mu$ g) was primed with the gene-specific primer MRP006 (5'-ACTTTACTCCGTAGTCATTAC-3') and reverse transcribed. After double-strand cDNA synthesis, the Marathon cDNA adaptor was ligated. The primary PCR reaction was performed using the adaptor-ligated cDNA and primers AP1 and MRP008 (5'-CTTTACCTTGGAATCGAGTGG-ACA-3'). For the nested PCR reaction, primers AP2 and MRP012 (5'-ACTGAATCCACCGGTCTCCATCCTT-3') were used. Amplified products were subcloned into the vector pT7-Blue(R) (Novagen) and their sequences were determined. We identified a clone, 5R#01, which contained the longest 1.0 kb insert. However, it lacked an initiation codon. To obtain fulllength coding sequences, we repeated the 5'-RACE as follows. For walking#1, primers AP1 and MRP022R (5'-CCCATTTT-CGTAGTGAATCTCCACAT-3') were used for the primary PCR reaction, and AP2 and MRP021R (5'-CCATTGTT-CAGTATGAGGTCAAGAATG-3') for nested PCR. For walking#2, primers AP1 and MRP026R (5'-TATTGTCTTC-ATCTGCCAGGACTCCC-3') were used for the primary PCR reaction and AP2 and MRP023R (5'-TTCTGCAGTTTCA-TTCCAATGCGGAG-3') for nested PCR. Amplified PCR fragments were subcloned to the pT7-Blue(R) vector (Novagen) and sequenced. We analysed more than 50 independent clones and identified the putative translation initiation codon. We constructed the human MDC3 protein coding sequences by connecting the overlapping cDNA sequences.

For the purpose of sequence verification, we cloned the cDNA by another method, that is, through screening of a human brain cDNA library. A lambda gt10 human brain cDNA library (Clontech) was plated according to manufacturer's instructions and screened with the MDC3 probe (5R#01) described above. Sixteen clones were isolated and each cDNA insert was amplified by PCR using lambda gt10 primers gt10Fw (5'-GACTGC-TGGGTAGTCCCCACCTTT-3') and gt10Rv (5'-TGGCTTA-TGATTTCTTCCAGGGTA-3'). Amplified products were purified and direct-sequenced using primer gt10Fw or gt10Rv. Out of 16 clones, nine clones were identified as MDC3, five clones as MDC/ADAM11, and two clones were classified as belonging to another novel MDC protein family (later termed MDC2). The sequences of the isolated clones from the cDNA library were identical with the putative human MDC3 protein coding sequences determined by PCR methods.

## **Cloning of MDC2 cDNA**

In the cDNA library screening of MDC3 described above, two clones containing about 1.0 kb of novel MDC-related sequences

were isolated. We have designated these cDNAs as MDC2, because they are closely related to MDC/ADAM11 and MDC3, but not identical with them. The EST database search with 1.0 kb human MDC3 partial cDNA sequences was performed, and it was shown that the mouse EST W96980 (Genbank accession number W96980) scored highly, with 88.7 % identity in 222 bp. We purchased the EST clone from Genome Systems Inc. (St. Louis, MO, U.S.A.) and determined its inserted sequences. The identified 900 bp of sequences revealed this clone to be potentially a murine counterpart of human MDC2, based on its high similarity to the human MDC2 cDNA. To obtain the human MDC2 cDNA, we performed reverse transcription (RT)-PCR using human brain poly(A)+ RNA and the primers MFP027(5'-CAGAGCATCATACCACTGCGCCTCATCT-3') and MFP010R (5'-ACATGAGATGATCATTCACAATCAT-CAG-3') designed from mouse EST sequences. We were able to identify 600 bp of novel sequences which overlapped with the 1.0 kb human MDC2 described above. To identify full-length coding sequences, 3'-RACE was performed as described above, except for templates (human and mouse) and primers MFP001 (5'-GAGGACACGTGGTCCGGGTGCATAATG-3') and UAP for primary PCR, MFP005 (5'-CCTTCTAAGCTTCTTGATC-CTCCTGAG-3') and UAP for nested PCR. Amplified products (1.2-1.6 kb) were subcloned into a pBluescriptII SK + vector (Stratagene) and their sequences were determined. Identification of the sequences from several 3'-RACE clones revealed the existence of two kinds of products from both the human and mouse. One was shorter than the other, 108 bp being deleted in the former. To identify the 5'-flanking sequences, we performed 5'-RACE using the 5'-RACE system. Human brain  $poly(A)^+$ RNA (1.0  $\mu$ g) was primed with MFP006 5'-CATCTTCCTCC-AAAGCAAATTCCCTGA-3' and reverse transcribed and dCtailed according to the manufacturer's protocol. For the primary PCR reaction, the primers AAP and MFP008R (5'-GTCAAG-GTGCATTTCTTACAACACTCT-3') were used and the primers UAP and MFP029R (5'-CAGCAAATCATGATTTA-GCACGACATC-3') were used for nested PCR. The amplified 5'-RACE fragments were subcloned into a pBluescriptII SK + vector (Stratagene) and their sequences were determined. The identified 500 bp sequence revealed the presence of a putative initiation codon with Kozak motif.

#### Northern blot analysis

Pre-made multiple tissue Northern blots (human MTN I, human brain MTN I,II) were purchased from Clontech. These membranes were hybridized with radiolabelled cDNA probes in ExpressHyb. solution (Clontech) according to the manufacturer's instructions. The cDNA fragments used for labelling were as follows: the human MDC/ADAM11 cDNA fragment containing 2.3 kb full-length coding sequences [9], the 0.5 kb human MDC2 cDNA fragment obtained by RT-PCR using primers MFP001 and MFP002 (5'-GTCTGATGCTGTCACCTTT-TGCCCCCCA-3') and human MDC3 probe 5R#01 described above.

#### **RESULTS AND DISCUSSION**

#### Cloning and sequencing of MDC2 and MDC3 cDNA

We have cloned two novel cDNAs from human brain. Analysis of their deduced amino acid sequences revealed that both possessed typical cellular disintegrin structures and exhibited high sequence similarity with human MDC/ADAM11 protein [8,9]. We therefore termed them MDC2 and MDC3 (Figure 1).

The predicted human MDC3 cDNA encodes an 832 amino

hum	MDC3	MKPPGSSSRQ	PPLAGCSLAG	ASCGPQRGPA	GSVPASAPAR	TPPCRLLLVL	50	
hum	MDC	LLLPPLAASS	MRLLRR	WAFAA <b>l</b> llsl	LPT <b>P</b> GLGTQG	PAGA <b>L</b> RWGGL	36	
hum	MDC2a		MQAAV <b>A</b> VS	VPFLL <b>LC</b> VLG	T <b>C</b> P <b>P</b> AR <b>C</b> GQA	GDAS <b>L</b> MELEK	38	
hum	MDC3		RPRAWGA <b>A</b> AP	SAPHWNETAE	KNLGVLADED	NTLQQNSSSN	100	
hum	MDC	PQLGGPGAP <b>E</b>	VTEPSRLV-R	ESSGG <b>EVR</b> KQ	QLDTRVRQEP	P <b>G</b> GPP <b>VHLAQ</b>	85	
hum	MDC2a	RKENRFVERQ	SIVPLRLIYR	SGGED <b>ESR</b> HD	ALDTRVRGDL	G <b>G</b> PQLT <b>H</b> VD <b>Q</b>	88	
hum	MDC3	ISYSNAMQK <b>E</b>	ITLPSRLIYY	INQDS <b>ES</b> PYH	VLDTKARHQQ	KHNKA <b>VHLAQ</b>	150	
hum	MDC	VSFVIPAFNS	NFTLDLELNH	H <b>llssqyver</b>	HFSREGTTQH	STGAGDHCYY	135	
hum	MDC2a	ASFQVDAFGT	SFILDVVLNH	D <b>llsseyier</b>	HIEHGGKTVE	VKG-GEHCYY	137	
hum	MDC3	ASFQIEAFGS	KFILDLILNN	G <b>llssdyve</b> I	HYEN-GKPQY	SKG-GEHCYY	198	
hum	MDC	QGKLRGNPHS	FAALSTCQGL	HGVFSDGNLT	YIVEPQEVAG	PWGAPQ <b>G</b> PL <b>P</b>	185	
hum	MDC2a	QGHIRGNPDS	FVALSTCHGL	HGMFYDGNHT	YLIEPEEND-	TTQEDFHF	184	
hum	MDC3	HGSIRGVKDS	KVALSTCNGL	HGMFEDDTFV	YMIEPLE-LV	HDEKST <b>G</b> R- <b>P</b>	246	
						→ Metalloprotainase	domain	
hum	MDC	HLIYRTPLLP	DP <b>L</b> G <b>C</b> REPG <b>C</b>	LFAVPAQ <b>S</b> AP	PNRP <b>RL-RRK</b>	RQVRRGHPTV	234	
hum	MDC2a	HSVYKSRLFE	F <b>SL</b> DDLPSEF	QQVNIT <b>PS</b> KF	I <b>L</b> KP <b>R</b> P <b>KR</b> S <b>K</b>	RQLRRYPRNV	234	
hum	MDC3	HIIQKTLAGQ	Y <b>S</b> KQMKNLTM	ERGDQW <b>P</b> FLS	E <b>L</b> QW- <b>LKRRK</b>	RAV-NPSRGI	294	
hum	MDC	HSETKYVELI	VINDHQLFEQ	MRQSVVLTSN	FAKSVVNLAD	VIYKEQLNTR	284	
hum	MDC2a	EEETKYIELM	IVNDHLMFKK	HRLSVVHTNT	YAKSVVNMAD	LIYKDQLKTR	284	
hum	MDC3	FEEMKYLELM	IVNDHKTYKK	HRSSHAHTNN	FAKSVVNLVD	SIYKEQLNTR	344	
hum	MDC	IVLVAMETWA	DG <b>DKI</b> QVQDD	LLETLARLMV	YRREGLPEPS	DATHLFSGRT	334	
hum	MDC2a	IVLVAMETWA	TDN <b>K</b> FAISEN	PLITLREFMK	YRRDFIKEKS	DAVHLFSGSQ	334	
hum	MDC3	VVLVAVETWT	EK <b>D</b> QIDITTN	PVQMLHEFSK	YR-QRIKQHA	DAVHLISRVT	393	
hum	MDC	FQSTSSGAAY	VGGICSLSHG	GGVNEYGNMG	AMAVTLAQ	TLG <b>QNLG</b> MMW	382	
hum	MDC2a	FESSRSGAAY	IGGICSLLKG	GGVNEFGKTD	LMAVTLAQ	SLAHNIGIIS	382	
hum	MDC3	FHYKRSSLSY	FGGVCSRTRG	VGVNEYGL	PMAVAQVLSQ	SLAQNLGIQW	441	
hum	MDC	NKHRSSA-GD	CKCPDIWLGC	IMEDTGFYLP	RKFSRCSIDE	YNQFLQEGGG	431	
hum	MDC2a	DK-RKLASGE	CKCEDTWSGC	IMGDTGYYLP	KKFTQCNIEE	YHDFLNSGGG	431	
hum	MDC3	EPSSRKPK	CDCTESWGGC	IMEETGVSHS	RKFSKCSILE	YRDFLQRGGG	489	
hum hum hum	MDC MDC2a MDC3	→ SCLFNKPLKL ACLFNKPSKL ACLFNRPTKL	disintegrin-like d LDPPECGNGF LDPPECGNGF FEPTECGNGY	omain VEAGEECDCG IETGEECDCG VEAGEECDCG	SVQECSRAGG TPAECVLEGA FHVECYGL	NCCKKCTLTH ECCKKCTLTQ -CCKKCSLSN	481 481 536	
hum hum hum	MDC MDC2a MDC3	DAMCSDGLCC DSQCSDGLCC GAHCSDGPCC	RRCKYEPR KKCKFQPM NNTSCLFQPR	GVSCREAVNE GTVCREAVND GYECRDAVNE +++++++	CDIAETCTGD CDIRETCSGN CDITEYCTGD ++	SSQCPPNLHK SSQCAPNIHK SGQCPPNLHK	529 529 586	
→ cystein-rich domain								
hum	MDC	LDGYYCDHEQ	GRCYGGRCKT	RDRQCQVLWG	H-AAA-DRFC	YEKLNVEGTE	577	
hum	MDC2a	MDGYSCDGVQ	GICFGGRCKT	RDRQCKYIWG	QKVTASDKYC	YEKLNIEGTE	579	
hum	MDC3	QDGYACNQNQ	GRCYNGECKT	RDNQCQYIWG	TKAAGSDKFC	YEKLNTEGTE	636	
hum	MDC	RGSCGRKGSG	WVQCSKQDVL	CGFLLCVNIS	GAPRLGDLVG	DISSVTFYHQ	627	
hum	MDC2a	KGNCGKDKDT	WIQCNKRDVL	CGYLLCTNIG	NIPRLGELDG	EITSTLVVQQ	629	
hum	MDC3	KGNCGKDGDR	WIQCSKHDVF	CGFLLCTNLT	RAPRIGQLQG	EIIPTSFYHQ	686	
hum	MDC	GKELDCRGGH	VQLADGSDLS	YVEDGTACGP	NMLCLDHRCL	PASAFNFSTC	677	
hum	MDC2a	GRTLNCSGGH	VKLEEDVDLG	YVEDGTPCGP	QMMCLEHRCL	PVASFNFSTC	679	
hum	MDC3	GRVIDCSGAH	VVLDDDTDVG	YVEDGTPCGP	SMMCLDRKCL	QIQALNMSSC	736	
hum hum hum	MDC MDC2a MDC3	→ E PGSGERRICS LSSKEGTICS PLDSKGKVCS	GF-like domain HHGVCSNEGK GNGVCSNELK GHGVCSNEAT	CICQPDWTGK CVCNRHWIGS CICDFTWAGT	DCSIHNPLPT DCNTYFPHND DCSIRDPVRN	SP <b>PTGETE</b> RY DAK <b>TGIT</b> LSG LH <b>P</b> PKD- <b>E</b> GP	727 729 785	
hum hum hum	MDC MDC2a MDC3	→ tra KGPSGTNIII NGVAGTNIII KGPSATNLII ***	nsmembrane dor GSIAGAVLVA GIIAGTILVL GSIAGAILVA *********	nain AIVLGGTGWG ALILGITAWG AIVLGGTGWG ******	→ cytoplasmic of FKNIRRGRSG YKNYREQRQL FKNVKKRRFD *	domain→→→ GA <u>PQGDYVKKPG</u> PTQQGPI	<b>769</b> 779 <b>832</b>	
hum	MDC2a	DGDSFYSDIP	PGVSTNSASS	SKKRSNGLSH	SWSERIPDTK	HISDICENGR	829	
hum	MDC2a	PRSNSWQGNL	GGNKKKIRGK	RFRPRSNSTE			859	

#### Figure 1 Deduced protein sequences of human MDC, MDC2 $\alpha$ , MDC3

Conserved residues are shown in bold and all cysteine residues are shown in white text. Transmembrane hydrophobic regions are highlighted as \*\*\*. Putative integrin binding loops are indicated as + + +. The 36 residues underlined were deleted in MDC2 $\beta$ , but are present in MDC2 $\alpha$  (indicated as MDC2a). EGF, epidermal growth factor.

acid protein including a hydrophobic transmembrane domain and eight potential N-linked glycosylation sites. A comparison of human MDC3 with other cellular disintegrins revealed a sequence similarity of 52.2 %/726 amino acids with human MDC/ ADAM11, 33.0 %/630 amino acids with mouse meltrin  $\alpha$ /ADAM12 [10], 30.9 %/669 amino acids with human MDC9 [11], 30.1 %/658 amino acids with human MDC15 [12] and less than 30 % with others.

In the case of human MDC2, we have identified two kinds of transcripts. One was 108 bp shorter than the other due to a deletion in the cytoplasmic domain. We have confirmed that the

deletion is caused by the skipping of a 108 bp single exon by genome sequence analysis (results not shown). We have designated the long-form as MDC2 $\alpha$  and the exon-skipped form as MDC2 $\beta$ . The deduced protein sequence of human MDC2 $\alpha$ contained 859 amino acid residues including a long cytoplasmic tail, and that of MDC2 $\beta$  contained 823 residues, of which 36 residues were deleted (Figure 1, underlined residues) in the cytoplasmic domain. Comparison of human MDC2 $\alpha$  with other cellular disintegrins revealed the following sequence similarity: 56.0 %/720 amino acids with human MDC/ADAM11, 51.1 %/726 amino acids with human MDC3, 32.4 %/675 amino

-
0

	TRIGRA	VTMTHEMGHNLGMHHDEDKCNCNTCIMS
	JARA	VIMAHEMGHNLGIHHDTGSCSCGDYPCIMG
	mus FERa	ALMA <b>HE</b> LGHNLGIQHD HPTCTCGPKHFCLMG
	mus MELa	VTLAHELGHNFGMNHDTLERGCSCRMAAEKGGCIMN
	mus MS2	STMAHELGHNLGMSHDEDIPGCYCPEPREGGGCIMT
	hum MDC9	SIVAHELGHNLGMNHDDGRDCSCGAKSCIMN
	hum MDC15	SSIAHELGHSLGLDHDLPGNSCPCPGPAPAKTCIME
	hum TACE	LVTT <b>HE</b> LG <b>H</b> NF <b>G</b> AE <b>H</b> DPDGLAE <mark>C</mark> APNEDQGGKYV <b>M</b> Y
	hum MADM	ITFAHEVGHNFGSPHDSGTECTPGESKNLGQKENGNYIMY
		HE**H**G**H
	mus FERb	IILVQLLSLSM <b>G</b> LAYDRRLNK <mark>C</mark> QCGVPVCVMN
	mus CYRI	IVLSQLLGINL <b>G</b> LAYDDVYN <mark>C</mark> FCPGST CIMN
	hum MDC	VTLAQTLGQNL <b>G</b> MMWNKHRSSAGD <mark>C</mark> KCPDIWLGCIME
	hum MDC2	VTLAQSLA <b>h</b> ni <b>g</b> iisdkrklasge <mark>c</mark> k <mark>c</mark> edtwsg <mark>c</mark> i <b>m</b> g
	hum MDC3	QVLSQSLAQNL <b>G</b> IQWE-PSSRKPKCDCTESWGGCIME
	В	
TRIGRA		SFIEEGTVCRIA <u>rgd</u> -dlddycngrsagc
	JARA C	KFSKSGTECRASMSECDPAEHCTGQSSEC
	mus FERa C	TFKKKGSLCRPAEDVCDLPEYCDGSTQEC
	mus MELa C	QLKPPGTACRGSSNSCDLPEFCTGTAPHC
	mus MS2 C	KVKPAGEVCRLSKDKCDLEEFCDGRKPTC
	hum MDC9 C	RFLPG <b>G</b> TL <b>CR</b> GKTSECDVPEYCNGSSQFC
	hum MDC15 C	QLRPSGWQCRPT <u>RGD</u> CDLPEFCPGDSSQC
	hum TACE C	QFETAQKKQQEAINATCKGVSY-C
	hum MADM C	AFKSKSEK <b>GR</b> DD-SD <b>G</b> AREGI <b>CNG</b> FTALC
		RGD
	mus FERb C	KLKRKGEV <b>GR</b> LAQDE <b>CD</b> VTEYCNGTSEVC
	mus CYRI C	TIAERGRL <mark>GR</mark> KSKDQ <b>CD</b> FPEF <b>C</b> NGETEGC
	hum MDC C	KYEPR <b>G</b> VS <b>CR</b> EAVNE <b>CD</b> IA <b>E</b> TCTGDSSQC
	hum MDC2 C	KFQPMGTVCREAVNDCDIRETCSGNSSQC
	hum MDC3 C	lfqpr <b>g</b> ye <b>gr</b> davne <b>gd</b> it <b>e</b> y <b>g</b> tgdsgq <b>c</b>
		AVNECD

\*\*\*D\*\*

#### Figure 2 Alignment of protein sequences from snake venom and cellular disintegrins

Aligned proteins are indicated as follows: snake venom trigramin precursor region (TRIGRA), jararhagin (JARA), and cellular disintegrins: mouse fertilin  $\alpha$ /ADAM1 (mus FERa), mouse meltrin  $\alpha$ /ADAM1 (mus MELa), mouse MS2/CD156/ADAM8 (mus MS2), human MDC9/ADAM9 (hum MDC9), human MDC15/ADAM15 (hum MDC15), human TACE/ADAM17 (hum TACE), human MADM/ADAM10 (hum MADM), mouse fertilin  $\beta$ /ADAM2 (mus FERb), mouse cyritestin/ADAM3 (mus CYRI), human MDC/ADAM11 (hum MDC), human MDC2 and human MDC3. (**A**) Alignment of metalloproteinase-like domain. Active-site amino acids for zinc-binding (HExxHxxGxxH) are highlighted in bold and all cysteine residues are shown in white text. (**B**) Alignment of putative integrin-binding loops in the disintegrin-like domain; conserved residues are highlighted in bold and all cysteine residues are conserved in three MDCs (MDC, MDC2 and MDC3).

acids with mouse meltrin  $\alpha$ , 29.1 %/724 amino acids with human MDC9, 31.0 %/686 amino acids with human MDC15, and lower scores with others.

These results indicated that MDC/ADAM11, MDC2 and MDC3 proteins are more closely related than other known cellular disintegrins.

# Sequence comparison of the human MDC2, MDC3 and other cellular or snake venom disintegrins

Like other cellular disintegrins, MDC2 and MDC3 have multiple domain structures including a pro-, a metalloproteinase-like, a disintegrin-like, a cysteine-rich, an epidermal growth factor-like, a transmembrane and a cytoplasmic domain.

Within the metalloproteinase-like domain, both MDC2 and MDC3 lack HEXXHXXGXXH active-site amino acids for zinc binding (Figure 2A), which is critical for the proteinase activity.

All active metalloproteinases, trigramin precursor [13], jararhagin [14], human TACE [5,6] and human MADM/ADAM10 [15], conserve the zinc-binding motif. On the other hand, mouse fertilin  $\beta$ /ADAM2 [16] and mouse cyritestin/ADAM3 [17], which have been shown to act as integrin ligands, are disrupted in this motif.

Amino acid sequences around the putative integrin-binding loop of snake venoms and cellular disintegrins are shown in Figure 2(B). All cellular disintegrins with the exception of TACE exhibit striking similarity with the snake venom disintegrin, trigramin and jararhagin. It has been shown that trigramin binds to platelet  $\alpha IIb/\beta 3$  integrin via its RGD sequences and thereby inhibits fibrinogen-dependent platelet aggregation [18]. The other, jararahagin, is a high-molecular-mass haemorrhagic MDC protein. In the disintegrin domain, jararhagin has ECD residues in the place of the RGD motif and recognizes  $\alpha 2/\beta 1$  integrin and specifically inhibits platelet-collagen adhesion [19]. In the group of cellular disintegrins, mouse fertilin  $\beta$ /ADAM2 is also an integrin ligand and interacts with  $\alpha 6/\beta 1$  integrin on the egg surface. Mouse cyritestin/ADAM3 is a potential integrin or ECM ligand, because both antibodies raised against its disintegrin-domain and a peptide mimetic of the putative integrinbinding loop efficiently blocked sperm-egg adhesion and fusion [17]. MDC/ADAM11, MDC2 and MDC3 have highly conserved sequences, AVN(E/D)CD, in the putative integrin binding loop. In the snake disintegrin kistrin, the conversion of an RGD motif to LDV dramatically changed the binding specificity from  $\alpha v/\beta 3$ ,  $\alpha 5/\beta 1$  to  $\alpha 4/\beta 1$  integrins [20].

These findings show that the sequences in this loop play a critical role in receptor–ligand adhesion specificity and suggest that MDC/ADAM11, MDC2 and MDC3 selectively bind to the same integrin or ECM via conserved AVN(E/D)CD sequences.

# mRNA expression profiles of the Human MDC/ADAM11, MDC2 and MDC3

For some members of cellular disintegrins, tissue-specificity of mRNA expression was precisely evaluated. For example, fertilin  $\beta$ /ADAM2, cyritestin/ADAM3 and ADAM6 transcripts are expressed specifically in testis. On the other hand, fertilin  $\alpha$ /ADAM1 and ADAM5 are expressed in various tissues [21]. We have analysed the mRNA distributions of MDC/ADAM11, MDC2 and MDC3 by Northern blot analysis (Figure 3). mRNAs of approximately 5.0 kb for MDC/ADAM11, 9.5 kb for MDC2, and several transcripts ranging in size from 3.5 to 7.0 kb were detected for MDC3. All three of these MDCs were highly expressed in the brain, but were detected only slightly or not at all in other tissues. Interestingly, mRNA expression patterns of the MDC/ADAM11 and MDC2 were nearly identical. Both types of transcripts were abundant in the cerebellum and scarcely expressed in the spinal cord, whereas the expression pattern of MDC3 was clearly different.

Cell-cell interactions are believed to play crucial roles in neural function and development. Cell adhesion molecules, integrin, immunoglobulin and cadherin families are widely expressed in the nervous system. We have identified two new potential cell adhesion molecules, MDC2 and MDC3. These novel MDCs and MDC/ADAM11 are highly expressed in the brain. These molecules should provide very useful tools for understanding the roles of the cellular disintegrins in neural function.

We gratefully acknowledge Dr. Yusuke Nakamura for critical reading of the manuscript and suggestions.



### Figure 3 Northern-blot analysis of human MDC, MDC2 and MDC3

Human MTN (lanes 1–8), human brain MTN (lanes 9–16) and brain MTN II (lanes 17–24) were purchased from Clontech. These membranes were hybridized with each cDNA probe radiolabelled according to the manufacturer's instructions. Each lane contains approx. 2  $\mu$ g of poly(A)<sup>+</sup> RNA from various human tissues as follows: lane 1, heart; lane 2, brain; lane 3, placenta; lane 4, lung; lane 5, liver; lane 6, skeletal muscle; lane 7, kidney; lane 8, pancreas; lane 9, amygdala; lane 10, caudate nucleus; lane 11, corpus callosum; lane 12, hippocampus; lane 13, hypothalamus; lane 14, substantia nigra; lane 15, subthalamic nucleus; lane 16, thalamus; lane 17, cerebellum; lane 18, cerebral cortex; lane 19, medulla; lane 20, spinal cord; lane 21, occipital pole; lane 22, frontal lobe; lane 23, temporal lobe; lane 24, putamen.

## REFERENCES

- 1 Wolfsberg, T. G., Primakoff, P., Myles, D. G. and White, J. M. (1995) J. Cell Biol. 131, 275–278
- 2 Blobel, C. P. (1997) Cell 90, 589-592
- 3 Pan, D. and Rubin, G. M. (1997) Cell 90, 271-280
- 4 Rosendahl, M. S., Ko, S. C., Long, D. L., Brewer, M. T., Rosenzweig, B., Hedl, E., Anderson, L., Pyle, S. M., Moreland, J., Meyers, M. A. et al. (1997) J. Biol. Chem. 272, 24588–24593
- 5 Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S. et al. (1997) Nature (London) 385, 729–733
- 6 Moss, M. L., Jin, S. L., Milla, M. E., Bickett, D. M., Burkhart, W., Carter, H. L., Chen, W. J., Clay, W. C., Didsbury, J. R., Hassler, D. et al. (1997) Nature (London) 385, 733–736
- 7 Almeida, E. A., Huovila, A. P., Sutherland, A. E., Stephens, L. E., Calarco, P. G., Shaw, L. M., Mercurio, A. M., Sonnenberg, A., Primakoff, P., Myles, D. G. and White, J. M. (1995) Cell 81, 1095–1104

- 8 Emi, M., Katagiri, T., Harada, Y., Saito, H., Inazawa, J., Ito, I., Kasumi, F. and Nakamura, Y. (1993) Nature Genet. 5, 151–157
- 9 Katagiri, T., Harada, Y., Emi, M. and Nakamura, Y. (1995) Cytogenet. Cell Genet. 68, 39-44
- 10 Yagami-Hiromasa, T., Sato, T., Kurisaki, T., Kamijo, K., Nabeshima, Y. and Fujisawa-Sehara, A. (1995) Nature (London) 377, 652–656
- 11 Weskamp, G., Kratzschmar, J., Reid, M. S. and Blobel, C. P. (1996) J. Cell Biol. 132, 717–726
- 12 Kratzschmar, J., Lum, L. and Blobel, C. P. (1996) J. Biol. Chem. 271, 4593-4596
- 13 Neeper, M. P. and Jacobson, M. A. (1990) Nucleic Acids Res. 25, 4255
- 14 Paine, M. J., Desmond, H. P., Theakston, R. D. and Crampton, J. M. (1992) J. Biol. Chem. 267, 22869–22876

Received 26 March 1998/19 May 1998; accepted 21 May 1998

- 15 Howard, L., Lu, X., Mitchell, S., Griffiths, S. and Glynn, P. (1996) Biochem. J. **317**, 45–50
- 16 Evans, J. P., Schultz, R. M. and Kopf, G. S. (1995) J. Cell Sci. 108, 3267-3278
- 17 Yuan, R., Primakoff, P. and Myles, D. G. (1997) J. Cell Biol. 137, 105-112
- 18 Dennis, M. S., Henzel, W. J., Pitti, R. M., Lipari, M. T., Napier, M. A., Deisher, T. A., Bunting, S. and Lazarus, R. A. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 2471–2475
- De Luca, M., Ward, C. M., Ohmori, K., Andrews, R. K. and Berndt, M. C. (1995) Biochem. Biophys. Res. Commun. **206**, 570–576
- 20 Tselepis, V. H., Green, L. J. and Humphries, M. J. (1997) J. Biol. Chem. 272, 21341–21348
- Wolfsberg, T. G., Straight, P. D., Gerena, R. L., Huovila, A. P., Primakoff, P., Myles, D. G. and White, J. M. (1995) Dev. Biol. **169**, 378–383