

Supporting Information for

**Recombinant production, characterization and
fiber spinning of an engineered short major
ampullate spidroin (MaSp1s)**

Christopher Thamm[#] and Thomas Scheibel^{ #,†,¶,§,‡,⊥}*

[#]Lehrstuhl Biomaterialien, Fakultät für Ingenieurwissenschaften, Universität Bayreuth,
Universitätsstraße 30, 95440 Bayreuth, Germany

[†]Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), Universität Bayreuth,
Universitätsstraße 30, 95440 Bayreuth, Germany

[¶]Bayerisches Polymerinstitut (BPI), Universitätsstraße 30, 95440 Bayreuth, Germany

[§]Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Universität Bayreuth,
Universitätsstraße 30, 95440 Bayreuth, Germany

[†]Institut für Bio-Makromoleküle (bio-mac), Universität Bayreuth, Universitätsstraße 30,
95440 Bayreuth, Germany

[⊥]Bayreuther Materialzentrum (BayMAT), Universität Bayreuth, Universitätsstraße 30, 95440
Bayreuth, Germany

To whom correspondence should be addressed: thomas.scheibel@bm.uni-bayreuth.de

N1L

10	20	30	40	50
GQANTPWSSK	ANADAFINSF	ISAASNTGSF	SQDQMEDMSL	IGNTLMAAMD
60	70	80	90	100
NMGGGRITPSK	LQALDMAFAS	SVAEIAASEG	GDLGVTTNAI	ADALTSAYQ
110	120	130	140	150
TTGVVNSRFI	SEIRSLIGMF	AQASANDVYA	SAGSGSGGGG	YGASSASAAS
160	170	180		
ASAAAPSGVA	YQAPAQAQIS	FTLRGQQPVS		

Module A

10	20	
GPYGPGASAA	AAAAGGYGPG	SGQQ

Module Q

10	20
GPGQQGPGQQ	GPGQQGPGQQ

NR3

10	20	30	40	50
GAASAAAVSVG	GYGPQSSSAP	VASAAASRLS	SPAASSRVSS	AVSSLVSSGP
60	70	80	90	100
TNQAALSNTI	SSVVSQVSAS	NPGLSGCDVL	VQALLEVVSA	LVSILGSSSI
110	120			
GQINYGASAQ	YTQMVGQSVA	QALAG		

Figure S1. Amino acid sequences of N1L(AQ)₁₂NR3 domains and motifs.

A

CmMaSp1s	1	---MTWTTARALSFLVVVICSQSLFALCQS--PWQSASMAESFMFTYFSEALGQSGAFTKEQIDDITIASSIKLGVDKMRSGKTS	80
AtMaSp2	1	---MNWSIRLALLGFVVLSTQTIVFSAGQATPWENSQLAESFISRFLRFIGSGAESPQNLDNMSSIGDTLKTAIEKMAQSRSKSS	82
EaMaSp1	1	---MSWTAKLALLLFLVAC-QGSSSLASHTPWNTPLAENFMNSRFMQLSSMPGFTASQLDNMSTIAQSIVS1QSLAAQGRTS	81
LgMaSp1	1	MIKMLNSTRAL-IPRVLCTQGLYVLGQANTPWSSKQNADAFISAFMTPQSQGAFSSDQIDDMSVISNT'1MAAMDNGM-GRIT	82
LhMaSp1	1	TTTMTWSTRLALSFLFVLTCTQSLYALAQANTPWSSSKANADAFINSFISAASNTGSFSQDQMEDMSLIGNT'1MAAMDNGM-GRIT	83
LhMaSp2	1	MTTMNWSTRLVLSILVVLCTQSLCALGQANTPWSSKENADAFIGAFMNAQSQSGAFSSDQIDDMSVISNT'1MAAMDNGM-GRIT	83
NiMaSp2	1	---MSWST-LALAIIAVLSTQCIFIAGQANTPWSDTATADAFIQNFLGAVGSGAFTPQDQLDDMSTVGDTIMSAMDKMARSNKSS	81
		* . :	
CmMaSp1c	81	GSKLQAMNIAFASAVAEIATTEGGEQTAEVKTKAVADALAFAFFQTKGAVININFINEIKNLISMFQAT--NT	151
AtMaSp2	83	KSKLQALNNAFMFASSMAEIAVAVEQQGLSLEAKTNIAISALSAFLETTGYVNQVFNEIKTLIMFIAQASSNE	155
EaMaSp1	82	PNKLQALNNAFMFASSMAEIAASEEGGSSLSTKTSSIASAMSNAFLQTTGVVNKFISEIERSLIMFAQASAND	154
LgMaSp1	83	PTKLQALDMAFASSVAAETAAVE--GQNIVTTNAISDALTSAFYQTTGVVNKFISEIERSLIMFAQASAND	153
LhMaSp1	84	PSKLQALDMAFASSVAAETAASE--GGDLGVTTNAIADALTSAFYQTTGVVNKFISEIERSLIMFAQASAND	154
LhMaSp2	84	QSKLQALDMAFASSVAAETAADV--GQNIVGAATNAISDALTSAFYQTTGVVNQFIFTGJSSLIGMFQAQVSGNE	154
NiMaSp2	82	KSKLQALNNAFMFASSMAEIAAAVEQGGQSMDVKTNAIANALDSAFYMTTGSTMNQFVNEMRSLSINMLSAAVNE	154
		*****;*****;*****. : . *.: . *.: ** . *.* . * . * . : * . * . * . * . * . * . * . * . * . * . * . :	

B

CmMaSp1s	1	-RLSSPPNAASRISNSVSYLLSNGGPSSSIALSNTINNAVSQISASNPGLSTCDILVQALLELISALITYILRSATIGQVNSSLAGQSAQMVGQSVYRAFF	98
AtMaSp2	1	SRLSSPGAAKSKVSSAVTSLVSSGGPITNSAALSN1TSNVSSQ1SSSNPGLSGCDVLVQALLE1VSALVH1LGSAN1QVNSSGVGRSAS1VGQS1NQAFS	99
LgMaSp1	1	SALAAPATSARISSHASTLLSN-GPTNPASISNVISNAVSQISSNPGYSSCDILVQALLELVTALLTIIGSSNVNDINYGSSGQYAQMVSQSVQNVFG	98
LhMaSp2	1	SALSSPPTHARISSHASTLLSS-GPTNAAALSNVINSNAVSQVSAAPGSSCDVLVQALLEIITALSIILDSSSVQVNNGSSGQYAQIVGQSMQQAMG	98
LhMaSp1	1	SALAAPATSARISSHASALLSN-GPTNPASISNVISNAVSQISSNPGASACDVLVQALLELVTALLTIIGSSNIGSVNYDSSGQYAQVVTQSVQNAFA-----	98
NiMaSp2	1	SRLASPDSGARVASAVSNLVSS-GPTSSAALSSVNAVSQIGASNPGLSGCDVLVQALLEIVSACVTILSSSIQVNNGAA-----	82
		. . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . : *.* . :	

Figure S2. (A) Sequence comparison of MaSp NRN and NRC (B) domains of different spider species. AtMaSp2: *Argiope trifasciata* MaSp2 (NRN sequence: DQ059136, NRC sequence: AF350266), CmMaSp1s: *Cyrtophora moluccensis* MaSp1s (AGQ04592), EaMaSp1: *Euprosthenops australis* MaSp1 (AM259067), LgMaSp1: *Latrodectus geometricus* MaSp1 (NRN sequence: DQ059133S1, NRC sequence: DQ059133S2), LhMaSp1: *Latrodectus hesperus* MaSp1 (EF595246), LhMaSp2: *L. hesperus* MaSp2 (EF595245), NiMaSp2: *Nepenthes inaurata madagascariensis* MaSp2 (NRN sequence: DQ059135, NRC sequence: AF350278). Consensus symbols indicate positions with fully conserved residues (*) and conservations between groups of strongly similar (.) and weakly similar (.) properties.

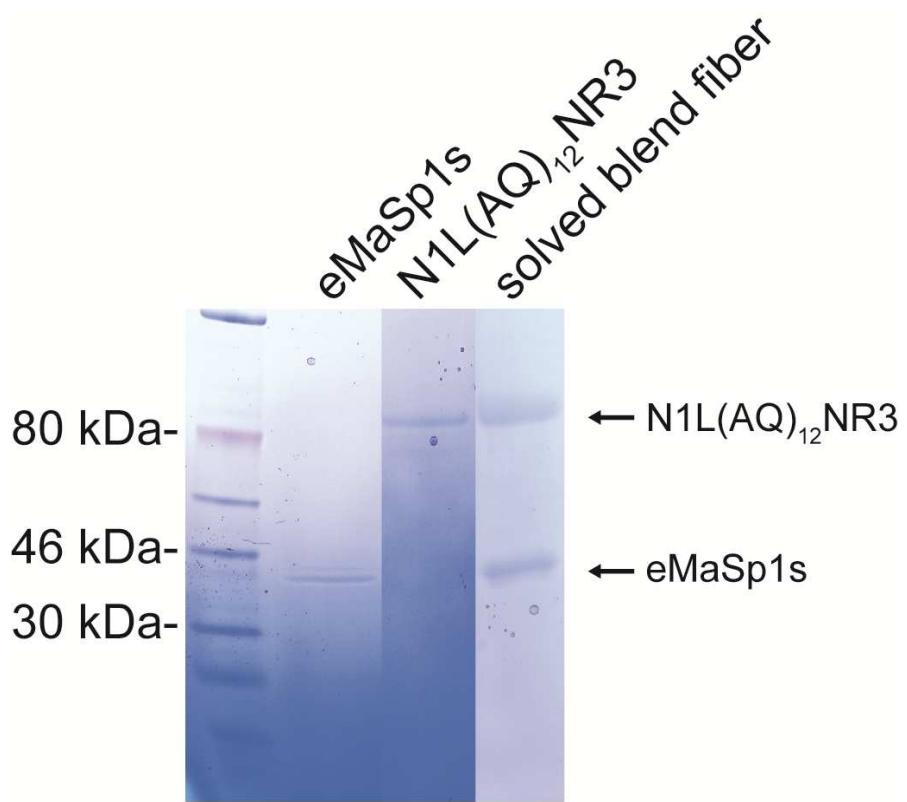


Figure S3. Coomassie stained SDS-gel analysis of eMaSp1s (purity >95%, determined using densitometry), N1L(AQ)₁₂NR3 and a resolved wet-spun blend fiber made of both proteins (eMaSp1s: 42.7 kDa, N1L(AQ)₁₂NR3: 87.7 kDa).

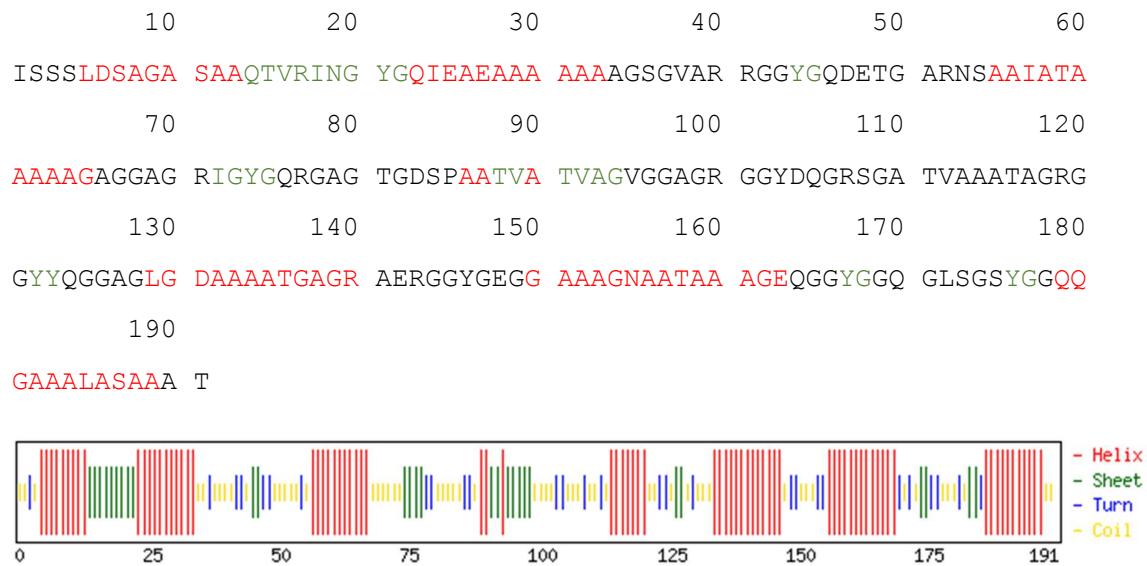


Figure S4. Sequence and secondary structure prediction of the eMaSp1s core domain. Prediction based on the method of Chou and Fasman,¹ typically used for analyzing soluble proteins.

REFERENCES

1. Chou, P. Y.; Fasman, G. D., Prediction of protein conformation. *Biochemistry* **1974**, *13* (2), 222-45.