Characterization of a Basidiomycota hydrophobin reveals the structural basis for a high-similarity Class I subdivision

Julie-Anne Gandier^{1,*}, David N. Langelaan^{2,*}, Amy Won³, Kylie O'Donnell¹, Julie L. Grondin⁴, Holly L. Spencer⁴, Philip Wong⁵, Elisabeth Tillier⁵, Christopher Yip³, Steven P. Smith⁴, Emma R. Master^{1,*}

¹Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario, Canada

²Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada ³Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

⁴Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada

⁵e-Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

^{*}These authors contributed equally to the work.

Correspondence and request for materials should be addressed to E.R.M. (email: emma.master@utoronto.ca)

Supplementary Materials

Completeness of Resonance Assignments (%) ^b	
Backbone	84.1
Side Chain	79.9
Aromatic	100
Number of conformational restraints	
Total NOE restraints	1495
Intraresidue (i=j)	647
Sequential $(i-j = 1)$	336
Medium range $(i-j < 5)$	167
Long range $(i-j \ge 5)$	345
Ambiguous	51
Dihedral angle restraints	132
Hydrogen bond restraints	50
NOE Restraints per residue	14.0
Long range restraints per residue	3.2
Residual restraint violations ^c	
RMSD NOE restraints (Å)	0.03 ± 0.001
RMSD dihedral angle restraints (°)	0.34 ± 0.052
Average no. of distance violations per structure	0
>0.5 Å	0
>0.3 Å	0.1
Average no. of dihedral angle violations per structure $> 5^{\circ}$	0.34 ± 0.05
Model quality ^d	
RMSD backbone atoms (Å) ^e	0.5
RMSD heavy atoms $(Å)^{e}$	0.8
RMSD bond lengths to ideal geometry (Å)	0.005
RMSD bond angles to ideal geometry (°)	0.8
MolProbity Ramachandran statistics ^{c,e}	
Most favoured regions (%)	92.2
Allowed regions (%)	7.4
Disallowed regions (%)	0.4
Global quality scores (raw/Z score) ^c	
Verify3D	0.38/-1.28
ProsaII	0.60/-0.21
PROCHECK $(\Phi - \Psi)^d$	-0.48/-1.57
PROCHECK (all) ^d	-0.63/-3.73
MolProbity clash score	37.12/-4.84
RPF analysis ^f	
Recall	0.981
Precision	0.729
F-measure	0.836
DP score	0.886
Model contents	
Ordered residue ranges ^e	10-116, 33-35, 37-81,91-115
Total no. of residues	107
BMRB accession number	25976
PDB ID code	2NBH

Supplementary Table 1: Structural statistics of the SC16 structural ensemble^a

^aStructural statistics computed for the ensemble of 20 deposited structures as recommended by the wwPDB task force.

^bCalculating using the AVS software. ^c Calculated using CNS version 1.21. ^d Calculated using the PSVS version 1.5. ^eBased on ordered residue ranges (S(f) +S(y) > 1.8). ^fCalculated using CcpNmr Analysis version 2.3.



Supplementary Figure 1. Sequence alignment of the core regions of SC3 and SC16.

The regions spanning from the first to last cysteines of SC3 and SC16 were aligned revealing 56% identity (ClustalW algorithm for pair-wise alignment with the Blosum cost matrix and free end gaps)



Supplementary Figure 2. Emission intensity vs. emission wavelength of Thioflavin T (ThT) excited at 442 nm . The maximum of the emission wavelength remained constant for solutions containing 35-66 μ g/ml aerated overnight by gentle tail over head shaking on a rotary shaker. N=1 for each curve.



Supplementary Figure 3. SC16 adopts a well-ordered structure. **A)** A 1 H- 15 N HSQC spectrum of SC16 with the assignments for backbone and side-chain resonances indicated. **B)** { 1 H}- 15 N NOE analysis of SC16 indicates that most of the protein is not dynamic on the ps-ns timescale.



Supplementary Figure 4. The electrostatic potential surface of SC16 has several charged patches. Negative and positive electrostatic surface potentials are coloured red and blue, respectively. Although there is a large white non-charged area (**D**), these residues are not hydrophobic in nature.



Supplementary Figure 5. SC16 and PcaHyd1 share similar secondary structures. The secondary structural elements of SC16 are indicated above the sequence alignment of SC16 and PcaHyd1. The chemical shift index based secondary structure prediction for SC16 (black) and PcaHyd1 (red) are aligned below the sequence alignment. The chemical shift index based prediction of secondary structure for SC16 and PcaHyd1 agree over well conserved regions of SC16 that are structured. Please note in the following stacks of Supplementary Figures 6A-AB all secretion signals were removed using SignalIP as describes in the methods of the main body of the paper. Tables of identities follow the alignments.

AP AP <th< th=""><th></th><th>1</th><th>10</th><th>20</th><th>30</th><th>40</th><th>50</th><th>60</th><th>70</th><th>80</th><th>90</th><th>100</th><th>110</th><th>120</th><th>131</th></th<>		1	10	20	30	40	50	60	70	80	90	100	110	120	131
Udentity 443 addresses bit with with with with with with with wi	Consensus	ΤP-			XX 🔣 X X X	KPASXCST	GPVQCCBXXX	XXAXXXAXX	KLGLLGIVX	QDLNVXVGLTC:	SPITVIGXXGS-	XCSAXXVCC	DNSXGGLIS	IGCVP	VXIXIXXXX
443 B LEV BG C = 2 B	Identity				and the second			a la companya da serie de la companya de							
447 Image: Control of the second	443	TP-			LEVRGGI	PPSSCST	GPVOCCNTV	TKASDPSATA	LELLGIVE	EDINVIVGLTC	SPITVIGVGGSN	ACSSNAVCC	ONNSFGELIS	IGCVP	VSL
941 942 944 945 946 947 948 948 948 948 948 948 948 948 948 948	447	TP-			LEVRGGI	SPPSSCST	GPVOCCNTV	TKASDPSATA	ILGULGIVI	EDINVVVGLTCS	PITVIGVGGSN	ACSSNAVCC	DNSNGGLIS	IGCVP	VSL
448 0	941	TP-			LEVRGGGGI	BPSSCST	GPVOCCNTV	TEASNPSATK	ILGLIGIVL	SDISVIVGLTCS	PITVIGVGGSN	ACTANAVCO	DNSNGGLIS	IGCVP	VSL
454 100 1	448	TP-			LEVRGGI	BPSSCST	GP I OC CNTV	TEASDPAAST	ILGLIGIVL	ODINVIVGLSCS	SPINVIGVGSSD	AGSANAVOO	NNSNGGLIS	IGCVP	VOL
944 02	454	TP-			LEVRGGI	PASSEST	GPVOCCNTV/	TANSASASA	ILGLIGIVL	SDINVPVGLDCS	SPINVIGVGGSN	A OSSNAVOO	DNSNGGLIS	IGCLP	VSL
852 10 10	944	TP-			LEVRGGI	PPSSCST	GPIOCCDTV/	TEASDPEATA	GLIGIVI	KDINVIVGLTCS	PITVIGVGGGS	AGSANAVE	DNSHGGLIS	IGCVP	VOL
114 112	852	TP-			LEGRGE	PASSCST	GPVOCCNTV/	TASDPAAAG	ILGLIGIVI	ODLTVAVGLSCS	SPISVIGVGTGN	AGSANAVOO	DNSLGGLIS	IGCVP	VTT.
518 ap	514	TP-			APRNE	PASSEST	GPVOCCGST	SASSSAAA	ALGILIGIVI	ODINVIVGLTCS	SPITVISASES-	SASAOAVCO	DNSHGGLIS	IGCVP	VTL
504 ab ab<	518	TP-			APRNE	PASSEST	GPVOCCDST	SASSSAAAI	GLIGIVI	ODINVIVGLTCS	SPITVIGASGS-	SASAOAVAA	DNSHGGLIS	IGCVP	VTI.
523 ab ab <t< td=""><td>504</td><td>TP-</td><td></td><td></td><td></td><td>PASSOST</td><td>GPVOCCDS</td><td>SASSSAAA</td><td>TELLGTVL</td><td>OD INVIVGL TCS</td><td>SPTTVTGASGS-</td><td>SASAOAVOO</td><td>DNSHGGLTS</td><td>IGCVP</td><td>VTT.</td></t<>	504	TP-				PASSOST	GPVOCCDS	SASSSAAA	TELLGTVL	OD INVIVGL TCS	SPTTVTGASGS-	SASAOAVOO	DNSHGGLTS	IGCVP	VTT.
519 a a b b c a b c b c b c b c b c	523	TP-			APRNE	PASSCST	GPVOCCDS	SASSSAAA	LELIGIVL	ODINVIVGLTCS	SPITVIGASCS-	SASAOAVA	DNSHGGLIS	IGCVP	VTL
527 p2	519	TP-			APRNE	PASSEST	GPVOCCDS	SASSSAAA	TELLGIVL	ODINVIVGLTCS	SPITVIGASGS-	SASAOAVCO	DNSOGGLIS	IGCVP	VTL
229 AP (2) =AP (2) =A	527	TP-			APRNE	PASSEST	GPVOCCDST	SASSSAAA	INGINGTVI	ODLNV VGLTCS	SPITVIGASGS-	SASMOAVAA	DNSHGGLIS	IGCVP	VTURT PGIS
513 13 <t< td=""><td>529</td><td>TP-</td><td></td><td></td><td></td><td>PASSOST</td><td>GPVOCCDS</td><td>SASSSAAA</td><td>INGINGTYL</td><td>ODINVIVGLTCS</td><td>PTTVTGASGS-</td><td>SASAOAVOO</td><td>DNSHGGLTS</td><td>IGCVP</td><td>TURTPGIS</td></t<>	529	TP-				PASSOST	GPVOCCDS	SASSSAAA	INGINGTYL	ODINVIVGLTCS	PTTVTGASGS-	SASAOAVOO	DNSHGGLTS	IGCVP	TURTPGIS
522 12	513	TP-			APRNE	PASSEST	GPVOCCDST	SASSSAAA	GUIGIVI	ODLNV VGLTCS	SPITVIGASGS-	SASAOAVAA	DNSHGGLIS	IGCVP	VTL
533 appendence appendence by the state of the	522	TP-			APCNE	PASSOST	GPVOCCDS	SASSSAAA	TELLGTVP	ODINVIVGLTCS	SPTTVTGASGS-	SASAOAVOO	DNSHGGLTS	IGCVP	VTT.
225 AP DINE	533	TP-			APCNE	PASSEST	GPVOCCDS	SASSSCA	TELLGTVP	ODTNVIVGT TC	PTTVTGASGS-	SASAOAVOO	DNSHGGTTS	IGCVP	VTT.
228 AP BY B	525	TP-				PASSOST	GPVOCCDS	SASSSAAA	TELLETVP	ODINVIVGLTCS	SPITVIGASCS-	SASAOAVOO	DNSHGGLTS	IGCVP	VTT.
330 AP BN E - AP BN E <t< td=""><td>528</td><td>TΡ.</td><td></td><td></td><td></td><td>PASSEST</td><td>GPVOCCDS</td><td>SASSSAAA</td><td>TIGHTGTVP</td><td>ODTNVIVGT TCS</td><td>PTTVTGASGS-</td><td>SASAOAVAG</td><td>DNSHGGTTS</td><td>IGCVP</td><td>VTT.</td></t<>	528	TΡ.				PASSEST	GPVOCCDS	SASSSAAA	TIGHTGTVP	ODTNVIVGT TCS	PTTVTGASGS-	SASAOAVAG	DNSHGGTTS	IGCVP	VTT.
APCGE	530	TP-			APRNE	PASSEST	GPVOCCDST	SASSSAAA	ILGILGTVP	ODINVIVGLTC	SPITVIGASGS-	SASAOAVOO	DNSHGGLIS	IGCVP	VTT.
S11 IP APRCE PASC SCREPVOCCOBTOTANSASATARG INCIVADENTAL OPTINUS (GRACSPANUIGAS CS - SCROAVCCE DNSHCOTSIGOTSIGOTS) S15 IP TPRNE PASC SCREPVOCCOSTOTANSASATARG INCIVADENTAL OPTINUS (GRACSPANUIGAS CS - SCROAVCCE DNSHCOTSIGOTS) S15 IP TPRNE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIGAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S26 IP TPRNE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIGAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S26 IP TPRNE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIGAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S26 IP TPRNE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIROCONNUCLICAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S26 IP TPRNE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIROS SASCA VICCONNUCLINGGOUND NUCLICAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S26 IP APRCE PASC SCREPTOCOSTO SASS SPULIAGOUND NUCLICOSTAVIROS SASCA VICCONNUCLICAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S01 IP APRCE PASC SCREPTICOSTO SASS SPULIAGOUND NUCLICAS CS - SCROAVCCE DNSHCOTSIGOTS VIA S01 IP IP </td <td>509</td> <td></td> <td></td> <td></td> <td>APCGE</td> <td>PASSEST</td> <td>GPVOCCDST</td> <td>TANSASATAI</td> <td>REGISTER</td> <td>ODTNV VGT TC</td> <td>SPTTVTGASGS-</td> <td>SASAOAVAG</td> <td>DNSHGGTTS</td> <td>IGCVP</td> <td>VTT.</td>	509				APCGE	PASSEST	GPVOCCDST	TANSASATAI	REGISTER	ODTNV VGT TC	SPTTVTGASGS-	SASAOAVAG	DNSHGGTTS	IGCVP	VTT.
515 AP TPRNE PASE OF DEPICEOS VORAS S PVAL ING INGIVIAD VIEN VERACES PLAYERAS (SES) SOLO AVCCEDNSICOPE	511	TP-			APRGE	PASSEST	GPVOCCDST	TANSASATA	I IGHIGTVI	ODINVIVGLTC	SPITVIGASGS-	SASAOAVOO	DNSHGGLIS	IGCVP	VTT.
535 AP TPRNE PASS CSRCPT QCC DSV0 SASS SPVAL INGINGIVER OVNEN VCHCSPTAVICAS GSLS SAA AVCCEDNSEGHTSTCCLP - Visit 526 AP	515	ΤP-			TPRNE	PASSOST	GPTOCCDSV	SASSEVAT	INGINGTVI	ODVNT NVGT TC	SPTTVTGASGS-	SASAOAVAG	DNSHGGLTS	IGCTP	VST
526 AP TPRNE ASS GERGET OCCOSING STS SPIVAL IAG INCIDENT NON VOLTE SPIANTICAS GERGET STGCHE - National State (State (Stat	535	ΠP-			TPENE	PASSOST	GPTOCCDSV	SASSEVET	TETTETVI	ODVNT.NVGT.TC	PTTVTGASGS-	SESAOAVOC	DNSHCGTTS	IGCTP	VST.
S01 AP APRGE PASE GROUP CONVERTENS OF ASTRONOMY PASES PRATING INCLUDING ODVERGED PARTING LESS - GONA WOOT DINSINGENTSTOCHE	526	ΤP-			TPRNE	PASSOST	GPTOCCDSV	STSSEV	INGINGTVI	ODVNV NVGL TC	SPTTVTGASGS-	SASAOAVAG	DNSHGGLTS	IGCTP	VSL
S10 AP AP <	501	TP-			APRGE	PASSEST	GDLOCCNTV	PASSPSAST	ILGUIGTVI	OGVDVIVGLTC	SPITVIGLESG-	GASAOAVCE	DNSNGGLIS	IGCTP	VTL.
750 (PcaHyd1) PAR PENAKRWATTTTPQAPTTTVTVTETATASSVPASD CSRCS VOCCAPVTSASD PMAG MG INGVV/OG VDVLVGHCD PLAVVGAG GG SSACAVCC DDNSY SSITSICCFPVIT 965 PPWG PTTPTVTVTVTAPASTATE	510	AD-			APRGE	PASSEST	GDLOCCNTV	PASSPSAST	GILGTVI	OGVDVIVGLTC	SPTTVTGLESG-	GASAOAVCO	DNSNGGLTS	TGCTP	VTT.
965 AP	750 (PcaHvd1)	TRN	PENAKRWATTT	TPOAPTTTVT	VTETATASSV	PASDOST	GSVOCCNTV	SASDPTAAG	ILGLIGVVV	OGVDVIVGLTC	PTTVVGAGGG-	SASAOAVAA	DNSYGSLIS	IGCTP	VTL.
621 AP TRETD QPASE ON GEVOCONDUTINAG DPANAG DEGINGTIVO DVN PVEDICAPIAVIGAG GS-GNAH PACEDNSEGTITISTCOVP 622 AP TRETD QPASE ON GEVOCONDUTINAG DPANAG DEGINGTIVO DVN PVEDICAPIAVIGAG GS-GNAH PACEDNSEGTITISTCOVP 622 AP TRETD QPASE ON GEVOCONDUTINAG DPANAG DEGINGTIVO DVN PVEDICAPIAVIGAG GS-GNAH PACEDNSEGTITISTCOVP 1070 AP TRETD TRETD TRETD VOCONDUTINAG DPANAG DPANAG DEGINGTIVO DVN PVEDICEPIAVIGAG GS-GNAH PACEDNSEGTITISTCOVP 626 AP TRETD TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVODINV PVEDICEPIAVIGAG GS-GSAS PVCCEDNSEGTISTGOVP VD DHU 634 AP TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVVODINV PVEDICEPIAVIGAG GS-GSAS PVCCEDNSEGTISTGOVP VD DHU 634 AP TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVVODINV PVEDICEPIAVIGAG GS-GSAS PVCCEDNSEGTISTGOVP VD DHU 637 AP TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVVODINV PVEDICEPIAVIGAG GS-GSAS PVCCEDNSEGTISTGOVP VD DHU 627 AP TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVVODINV PVEDICEPIAVICAG GS-GSAS PVCCEDNSEGTISTGOVP VD DEME 625 AP TRETD KPASE OST GEVOCONDUTINAG DPANAG VIA LIGVVVODINV PVEDICEPIAVICAG GS-GSAS PVCCEDNSEGTISTGOVP VD DEME	965	TR-	WGOPT	TPTVTVTV	TAPASTATE	PASNEST	GPIOCCDSVI	TASDPAVAT	LEBLIGIVV	EGLDVILGINCS	SPVSVIGVGGG-	SASAEAVCC	DNSLGSLIS	IGCVP	VTL
622 AP TRETD CPASE ON GEVOCONDUTING DPANG DIG INGELVODUNT PUBLICSPIRULGAGES - GUNAP VICEDNSIG 1070 AP TRETD CPASE ON GEVOCONDUTING DPANG DIG INGELVODUNT PUBLICSPIRULGAGES - GUNAP VICEDNSIG 1070 AP TRETD CPASE ON GEVOCONDUTING DPANG DIG INGELVODUNT PUBLICSPIRULGAGES - GUNAP VICEDNSIG 626 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA ING DPANG VIA DINV PUBLICSPIRULGAGES - GUNAP VICEDNSIG 696 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPPUAVICAGES - GUNAP VICEDNSIG 634 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPVAVIGAGES - GUNAP VICEDNSIGOUSSIGOVP - VIDINI 634 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPVAVIGAGES - GUNAP VICEDNSIGOUSSIGOVP - VIDINI 637 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPVAVIGAGES - GUNAP VICEDNSIGOUSSIGOVP - VIDINI 637 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPVAVIGAGES - GUNAP VICEDNSIGOUSSIGOVP - VIDINI 627 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DIAVEVOLACEPVAVIGAGES - GUNAP VICE DISSIGOUSSIGOVP - VIDINI 628 AP TRETD KPASE OST GEVOCONDUTING DPANG VIA DINVEVIDACEPVAVIGAGES - GUNAP VICE DISSIGOUSSIGOVP	621	TP-			TRRTD(PASECNT	GPVOCCNTV/	INAGDPARAG	TGLIGTIV	ODVNIPVGLTC	PTTVTGAGGS-	GONAHPAG	DNSHGTLIS	IGCVP	VSIGL
1070 1070	622	TP-			TRETD(PASECNT	GPVOCCNTV/	INAGDPAAAG	TGLIGTTV	ODVNT PVGLTCS	PTTVTGAGGS-	GANAHPMAA	DNSHGTTTS	IGCVP	VSTGL
626 AP	1070	TP-			TRRTDI	RASECTT	GPVOCCDTV/	TNAGDPSAAK	TGLIGIVV	ODINVPVGLTCS	SPITVIGAGGS-	GASAHPWAA	NNSHGULIS	IGCVP	VDLNL
696 AP	626	TP-			TRRTDI	PASECST	GPVOCCNTT	NAGDPTAAG	TALIGVVV	ODINVPVGLTC'	PVTVTGAGGS-	GASASPICA	DNSNGGLTS	IGCVP	VDI.HI.
634 DE	696	ΠP-			TRETD1	PASECST	GPVOCCNTT	INAGDETAAG	TALLGVVV	ODT.NVPVGT.TC	PVTVTGAGGS-	GUSASPICE	DNSNGGLTS	IGCVP	VDT NT.
627 AP	634	TP-			TRRTDI	PASECST	GPVOCCNTT	NAGDPTAAG	TALIGVVV	ODINVPVGLTC'	PVTVTGAGGS-	GASASPICA	DNSNGGLTS	IGCVP	VDLKL
625 AP	627	TP-			TRETGI	RASECST	GPVOCCNTT	INAGDPTAAG	LALIGVVV	ODINVPVGLTC	PVTVIGAGGS-	GASASPVCC	DNSNGGLIS	IGCVP	VELKL
630 AP	625	TP-			TRRTD1	PASECST	GPVOCCNTT	INAGDETAAG	TALIGVVV	ODINVPVGLTC	PVTVTGAGGS-	GASASPVOO	DNSNGGLTS	IGCVPSTS	TSKRIYFFL
	630	TP-			TREGGI	PASSEST	GP TOCCNTV	TNAGDKTAAG	LALIGVVV	ODINTIVGTTCS	SPITVIGAGGS-	GASANPWCC	DNSHGGLIS	IGCVP	VDIKI
	702	TR-			ARRGGG-1	PASSEST	APVOCCNTV	INAGDKTAAG	LALIGVVV	ODINTIVGITCS	SPITVIGAGGS-	GESANPVCC	DNSHGGLIS	IGCVP	VDLNL

Supplementary Figure 6A. Sequence alignment and corresponding consensus sequence of proteins contained in Region 1 (full sequence)

	1	10	20	30	40	50	60	70	80 9	90
Consensus	CSTGPVC	CCBXX	XXAXXXXAXXA	GLLGIVXQDLN	VXVGLTCSPI	VIGXXGS-X	CSAXXVCCX	DNSXGGLISI	GCVPVXIXXXXX	×.
Identity			المحالية							_
F10	a smanua									-
516	CSTGPVC	CCUST	CASSSSAAAF	GLLGIVLODIN	VEVGLICSPI			DNSHGGLISI		
514	COTGPV	CCBC		GLLGIVDODIN	VEVGLICSPI			DNSHGGLISI		
522	CSTGPVC	CCDS		CLIGIVPODIN	VIVCLTCSPI			DNSHCGLISI		
533	CSTCPVC	CCDS	S S A S S S S C C A VI	CLIGIVPODIN CLICIVPODIN	VINCLUCSPI			DNSHCGLIST		
530	CSTGPVC	CCDST	SASSSSAATT	CITCTVPODTN	VINCTTCSPT	VIGASCS-S	SCSAOAVCCD	DNSHGGTTST	GCVPWITT	
528	CSTGPVC	CCDST	SASSSSAATT	GLIGIVPODIN	VINCLTCSPT	VTGASGS-	SCSAOAVCCD	DNSHGGLIST	GCVPWTT.	
529	CSTGPVC	CCDST	SASSSSAATT	GLIGIVLODIN	VIVGLTCSPIT	VICASES-	SCSAOAVCCD	DNSHGGLISI	CVPVTIRTPGI	s
527	CSTGPVC	CCDST	SASSSSAAATT	GLLGIVLODIN	VIVGLTCSPIT	VICASES-	SCSAOAVCCD	DNSHGGLISI	CVPVTIRTPGI	s
513	CSTGPVC	CCDSI	SASSSSAAAII	GLLGIVLÖDLN	VIVGLTCSPIT	VIGASGS-S	SCSAOAVCCD	DNSHGGLISI	GCVPVTL	
523	CSTGPVC	CCDS	SASSSSAAAVI	GLLGIVLÕDLN	VIVGLTCSPIT	VIGASGS-	SCSAOAVCCD	DNSHGGLISI	GCVPVIL	
504	CSTGPVC	CCDS	SASSSSAAAVI	GLLGIVLQDLN	VIVGLTCSPIT	VIGASGS-	SCSAQAVCCD	DNSHGGLISI	GCVPVTL	
519	CSTGPVC	CCDS	SASSSSAAAVI	GLLGIVLQDLN	VIVGLTCSPIT	VIGASGS-	SCSAQAVCCD	DNSOGGLISI	GCVPVTL	
511	CSTGPVC	CCDS	QTANSA SA TALI	GLLGIVLQDLN	VIVGLTCSPIT	VIGASGS-S	SCSAQAVCCE	DNSHGGLISI	GCVPVTL	
509	CSTGPVC	CCDS	QTANSA SA TALI	GLLGIVLQDLN	VIVGLTCSPIT	VIGASGS-S	SCSAQAVCCE	DNSHGGLISI	GCVPVIL	
535	CSTGPIC	CCDSV	OSASSSPVAILI	GLLGIVLQDVN	LNVGLTCSPIT	VIGASGS-	SCSAQAVCCE	DNSHGGLISI	GCLPVSL	
515	CSTGPIC	CCDSV	QSASSSPVAILL	GLLGIVLQDVN	LNVGLTCSPIT	VIGASGS-	SCSAQAVCCE	DNSHGGLISI	GCLPVSL	
526	CSTGPIC	CCDSV	QSTSSSPVAIL	GLLGIVLQDVN	VNVGLTCSPIT	VIGASGS-S	SCSAQAVCCE	DNSHGGLISI	GCLPVSL	
510	CSTGDLC	CCNTV	EPASSPSASTII	GLTCIAIÓGAD	VIVGLICSPI	VIGLESG-0	GCSAQAVCCT	DNSNGGLISI	GCLPVIL	
501	CSTGDLC	CCNTV	EPASSPSASTII	GTTGIAIÖGAD	VIVGLTCSPIT	VICLESG-C	GCSAQAVCCT	DNSNGGLISI	GCLPVTL	
750 (PcaHyd1)	CSTGSVC	CCNTV	TSASDPTAAGIII	GTTGAAAOGAD	VIVGLICDPI		SCSAQAVCCQ	DNSYGSLISI	GCIPVIL	
965	CSTGPIC	CCDSV	ETAS DPAVATU	BUTCIANEGUD	VILGINCSPVS	VIEVGEG-S	SCSAEAVCCE	DNSLGSLISI	GCVPVIL	
447	CSTGPVC	CCNTV	TKASDPSATAL	GLLGIVLEDLN	VEVGLTCSPIT	VIGVGGSN	ACSSNAVCCO	DNSNGGLISI	GCVPVSL	
443	CSTGPVC	CCNTV	TKASDPSATAL	GLLGIVLEDLN	VEVGLICSPIT	VIGVGGSN	ACSSNAVCCO	NNSFGGLISI	GCVPVSL	
941	CSTGPVC	CONT	TEASNPSATKI	GLIGIVISDIS	VIVGLICSPIN			DNSNGGLISI		
852	CSTGPV	CCNTV		GLIGIVISDIN		VIGVGG5NI				
448	CSTGPVC	CCNTT		CLIGIVIODIN	VIVCLECEDI			NNSNCGLIST	COVPWOI	
944	CSTCPTC		TEASDEAGEI	CLIGIVIQUIN	VINCLUCSPI	WTEVCECS		DNSHCGLIST		
622	CNTGPVC	CCNTV	TNACDPAAACT	CLIGITVODVN	TPVGLTCSPT	VIEACES-(CNAHPVCCR	DNSHGUTTST		
621	CNTGPVC	CCNTV	INAGDPAAAGT	GLIGITVODVN	TPVGLTCTPT	VIGAGES-(CNAHPACCE	DNSHGTLISI	GCVPVSTGT	
1070	CTTGPVC	CCDTV	TNAGDPSAAKT	GLLGIVVODIN	VPVGLTCSPIT	VICAGES-	GCSAHPWCCE	NNSHGTLISI	GCVPVDI NI	
702	CSTAPVC	CCNTV	TNAGDKTAAGU	ALIGVVVODIN	TINGITCSPIT	VICAGES-	CSANPVCCS	DNSHGGLISI	GCVPVDI NI	
630	CSTGPIC	CCNTV	TNAGDKTAAGII	ALIGVVVODLN	INGITCSPIT	VIGAGGS-	GCSANPVCCS	DNSHGGLISI	GCVPVDIKI	
696	CSTGPVC	CCNTT	INAGDP TAAGVI	ALIGVVVQDLN	VPVGLTCTPVT	VIGAGGS-(GCSASPVCCE	DNSNGGLISI	GCVPVDINI	
626	CSTGPVC	CCNTT	INAGDP TAAGVI	ALIGVVVQDLN	VPVGLTCTPVT	VIGAGES-0	GCSASPVCCE	DNSNGGLISI	GCVPVDIHI	
634	CSTGPVC	CCNTI	INAGDP TAAGVI	ALIGVVVQDLN	VPVGLTCTPV	VIGAGGS-(GCSASPVCCE	DNSNGGLISI	GCVPVDIKI	
627	CSTGPVC	CCNTT	INAGDP TAAGVI	ALIGVVVQDLN	VPVGLTCTPV	VIGAGGS-(GCSASPVCCE	DNSNGGLISI	GCVPVELKI	
625	CSTGPVC	CCNTI	INAGDP TAAGVI	ALIGVVVQDLN	VPVGLTCTPV	VIGAGGS-(GCSASPVCCE	DNSNGGLISI	GCVPSTSTSKRUYFFI	L

Supplementary Figure 6B. Sequence alignment and corresponding consensus sequence of proteins contained in Region 1 (without N terminus region preceding first cysteine of the eight-cysteine pattern)

C	1 10	20	30	40	50	60	70	80	90	100	110	120	127
Identity	APGTEXEXXXXXX				exxoleexx	WXX <u>M</u> X <u>S</u> XXXXX			O SIE II K X X E K EP	GXXGXXXX		LISIGCXPXXI	XLK
814		C D U P	TRADOR	0.55						own Delbus		TRICCIPUT	
639		TPVE.	LEAROS		CSTOCCES		THE REPORT	Chington	CSPLIVVGVGS	CNA SPINA		TISTCOUPUT	3
848		TPVR.	VENDASS		GSTOCCSS	VOSSDBASSNI	TRADETVI	GDUTGLICIC	CSPTTVVGVG	GNA SENA		TTSTCCTPTR	á –
902		TPLE	ARODPTSO	CNT	dstoccso	TESASSESASE	FIGL REIVE	EDUTALICIO	CSPISVIGVG	GSACTASP	VCCSNTAVGG	LIGIGCVPISI	1
930		TPSP	IVTRDGS D	CST	GP TOCCOO	TATEGSAAGAA	ALGLVGVVI	EDASVLIGIC	CSPISVIGVG	GSACDASP	VCCENNSVGG	LISIGCVPVAL	á –
942	MPGG	OPPSTTTRT	ITVTOTATG	PEPTSSTT	GPIOCCOS	EKPTDHDAS	TINGLIND TWO	EGPDVLVGI	CSPISVIGVG	GNSOSANV	VCCONNNVGG	LISIGCVPIII	í i
849	KPWN	TTPPPVTTTAT.	ATVTVTVSAL	PAPTGGD-TCST	GPIOCCNS	VGSASDPAFSO	INGLUCIVE	EGVEALVGL	CSPISVVGVG	GNACSATT	VCCONNSVGG	LISIGCIPIII	i i
964	N 2 WG	APP-TTTSTPT	VTVTVTAPAS	STVTAAS-OCST	GSVOCCES	TESANSGLGGI	FIGLEGIV	EGVDVLLGVN	CSPISVIGVG	GN-CNSNV	VCCODNAVGG	LISIGCIPII	
893	SPWD	TPTTTATKT	VTETVTAPAR	PTVTTAG-QCNT	GDAOCCOS	TESASSAAGS	FIGL RELVE	EDADVL LGL	CSPLDVVGVG	GSECNAEP	VCCTENGIGD	LISIGCVPISI	ā i
438	PNEONAKRWD	VSTTTATAT	ATVTVTA	-TPSATP-SCST	GPVQCCNS	LEKAS SPAGSI	LIGUV	QGVDVL IGL D	CSPISVVAVG	GQ-CKNNV	VCCONNAVGG	LISVGCIPVII	
811	TPME	VKRGGAPPPVT	KTVTVTA PA	ATPTGGSCST	GPIQCCNS	VQSSKSDAGSI	LIGIUG	QGVDVL LGV G	CSPISVIGVG	GS-ONNT	VCC ONNNVGG	LISIGCVPIII	
933	PPPPITLD	VRGGEPTVT	KTITVTAPGI	PTKTITAGDCT	GPIQCCNS	VQEASSQEVIS	ILGIUGIVI	NDETALVGLO	CSPISIIGI VO	GGACSBOP	VCCENNSVNS	IISIGCIPIFI	
931	PNPN-MKRWG	SPTTTPPVT	KTVTVTAPGI	PTTTVSAGNCNT	GDIQCCNT	VTKASSGEASS	ILGLES IVV	SDITTLVGL	CSPLSVIGIG	GSACSAHP	VCCONNSEGG	LISIGCIPIII	
899	PRPNPDNAKRWG	VPTTSATVT	TTITVTAPAR	PTVTTVS-QCNT	AP IQCC EQ	VEGDSDAATA	ILGLUGIVI	TDIDVLIGIN	CSPITIIGVG	GNECSATP	VCCENNSVGG	LISIGCIPIII	4
753	PNPIAEKRGG	QPS	QTVTVTAPAT	TTPTSVS-TCST	GDAQCCNS	VISASNPAAAG	ILSG LEIVI	OG FAD VGL	CSPITAVGAG	G-SCDSQA	VCC DNS GG	LISLGCLPITI	
751	DPNPENAKRWG	SPPATT	TTVTVTAPAS	STPTSVS-T <mark>CST</mark>	GGAQCCNS	VISISDPAIS	LIGIUU	QGIGAT VGL D	CSPINVGV30	G-TCDQQA	VCC DNSNS-	LVSIGCEPIQI	
842	APGT NAERMARGLP	PKAPRKLYRKD	PTPASPARRS	STPSGSPSSCST	GPVQCCNT	LTTASDPLAS	ILGLEGIII	G-PDVA VGL I	CSPISVIGVG	GS-CSANT	VCCENNNVGG	LISIGCIPIII	
822		S	PMPNTPA	PQSCST	GPIQCCNQ	VISASSSLASV	ILGALGLVV	OD FNLP VGI	CAPIT GIGLG	GSACSAIT	VCCEDNSVGS	LISIGCIPIII	
884	MPWE	TTSPATST	TTVTVTAPAS	STSTEPASSCDT	GSLQCCDT	VTS <mark>A</mark> KSESASA	ILGILGVVI	SNEDVLVGL	CSPISVISSG	SG-CTADP	VCCENNNESG	LISLGCVPVNI	SLK
828			TAVG	ATAVIPRSTCST	geikccns	VQQANSPSAAA	LLGLUGVVV	QGVDVLVGL	CDPISVVGAG	N S - CS N T	VCCENNSYGG	LISIGCVPVEI	4
1037	TPYPN	GGGGGGGGNPPA	TTTVTVTAP	GPTVTAVSQCNT	GDVQCCNT	VETADGPTAAG	LLALLGVVV	ODINVLVGI	CSPITILGIG	G-ECSARP	VCCSNNSFGG	LISIGCVPISI	L
815	LP-	GGKPPPP	TTTITATGPI	PPTSTPGDLCST	GNVQCCQS	VGGADDHVIGA	ELGVV	DAVDVVLGL	CSPIKIVGLG	DKACSSNV	VCCENNSIGG	LISIGCIPIII	
834		TPQ.	A	AGSCST	GTLOCCNS	VEARGSAAVAR	ILAA LGVVV	OD INIP VGL S	CSGVNGVGVG	SDSCSTNA	VCCDENNFGG	LLAIGCEPASI	4
821		APNNTPT.	APAAPAPSAI	PAGGAGAGS	GALOCCOS	VESGSSAAVAR	ILAA LGVVI	QDVNIP VGL S	CSDITEVGVG	GNSCSMNT	VCCDDNSFGG	LLSVGCEPASI	
830		TPA	P	AADASCST	GALOCCOS	VEPAGSEAVAE	ILLAG LGVVV	OD NIPVGVI	CRAITVVGVGS	GSSCSENT	VCCDDNSFGG	LVSIGCEPAS	4
833		TPA	P	AADAS	GAIQCCDS	VOKADSAAVAF	ILAG LGVVV	QD VNAL VGL	CSPISVVGVG	SDACSMNA	VCC ODNSEGG	LVSIGCAPVSI	4
812		SPT	P	QAGSCST	GALOCCNS	VOSADSLGIAF	INDAA II GVIII	QDINIPIGVI	CSPISVVGVG	GNACSENA	CCENNSIGG	LISIGCEPAII	4
951		APA.	AP	PASSCST	GPIQCCES	VOPMGSAAATS	DEASIGVVV	ODVETPIGI	CSPISVVNVG	SDACSMOT	VCCEDNAFGG	LVSIGCIPVSV	4
442		TPN	P	VVARGESCST	GPVQCCNS	VOKEDSPAAAC	INKSIGVVV	ODVDALVGL	CSPISVVGVG	GSTCSGTT	VCCENNSFGS	LISIGCVPIVI	4

Supplementary Figure 6C. Sequence alignment and corresponding consensus sequence of proteins contained in Region 2 (full sequence)

	1	10	20	30	40		50	60	70	ł	30 84
Consensus	CXTEXX	X X D D O X	XXXXXXXXXXXX	JLXXLEXVXXX	XXXXXXGLX	CSPIXXX	xexexx	CXXXXVCC	XBNXXGGLIS	IGCXPX	XLXLK
Identity											_
1037	CNTCD	TOCCMENT	ETA DO DTA A C		TNUTWET		TCOC-E		SNN SPCCTTS	TCCUPT	G TT.
884		OCCDTV	TCAVEFEACA	TTCLTCVVTQN		CCDICULO				CCVPT	NTCTZ
753	COTOD	OCCNEV	TGAGNDAAAC		MTA DVCT T	CSPISVIS	12 ACA_C		ODN SPCCT. TS		
751	CSTCCZ	OCCNSV	T S T S T F A A A G			CSPINVVC				TGCTPT	01
849	CSTGS		CGAGDDAFCC		MEA IVELC	CGPTGVVG	WAGANA			TGCTPT	
064	Conce		G SASDPAL SG		VEALVGLG	CEPTEVIC			ODNAV CCT.TC	TCCTPT	
428	COTOR		EVACEDACET		VDVIIGVN	CEPTEVV				IGCIPI	T 7
811	CGTCP		OSEKSDACST		VDV LTCVC	CSPISVIC			ONNNY CCT.TS	TCCVPT	TT
942			EXDTOTICS EV	TTCLTDIVTEC	TDV LVCL	CGPIGVIC	WAGANG	CGANVY CC	ONNNY CCT.TS	TCCVPT	TT
842	Camepa		TTASDULA SV		DVAVCT	CSPISVIC	Weces			TCCTPT	TT
902	CNTCC		F GAG G F G A G F			CGPTGVTC	WAGAGA		SNTAV CCT.TC	TCCVPT	C T
803	CNTCDZ		ESASSA ACSO		ADVITCID	CSPLDWVC	WASASE	CNAEDVCC	TENCT ODT.TS	TCCVPT	ST.
930	CGTCD		ATACSAACAA	ATCLUCVVTED	ASVITCIC	CGDTGVTC	WCGCGA	CDASPVCC	INNSV CCT. TS	ICCV PV	TT.
848	Camela		OGO GA GGNT	TTGLTGTVICD	TTC LICE	CSPITVVC	WESENA	CSANA 7 CC	INNNY CCLTS	TCCT.PT	RT.
814	CSTCA		OSESSVSCNF		WTG LICED	CSPLTVVG	Wesena	CSANAVCC		TGCLPV	TT
639	Comco	OCCNTV	EDAKSASASI		TGLICIN	CSPLNWVC	WESENA	CSANAVCC	ONNNYCCLTS	TCCVPV	TT
033			OFACCOFVES	TTCLTCTVT ND		CGDTGTTC	TVCCCA	CSA OPVCC		TCCTPT	ET.
933	CNTCD		TKASSCEASS	TTCLTSTVVSD	TTTTT	CSPISVIC	TACASA	CSAH DV CC	ONNSECCTTS	TCCTPT	TT
800	CNTAD		VECDSDAATA		UDV LUCIN	CSPTTTC	VGTONE	CSATDVCC		TGCTPT	TT
822	CSTOP		TSASSSIASV	LICALCIVY OD	NT. PVCTA	CAPTTETC	TesesA	CSATTVCC	DNSV GST.TS	TGCTPT	TT
815	CSTEN	ZOCCOSV	GGADDHVTGA	LIGI. ICVVV DA	VDV VI.CT.O	CSPTKTVC	LOTDKA	CSSNVVCC	DINSTRATS	TGCTPT	TT
951	CSTGP	TOCCESV	OPACSAAATS	ITA STEWAYOD	Ψ ΨΨΡ ΤC ΤΨ	CSPTSVVN	WEGSDA	CSADTVCC	EDNAEGGLVS	TGCTPV	57
833	CSTGA	TOCCDSV	OKADSAAVAP	TTAGTGVVVOD	VNA LVCLT	CSPISVVG	VEGSDA	CSANAVCC	ODNSEGGIVS	TGCAPV	ST
834	CSTGT	OCCNSV	EAAGSAAVAP	TTA A LEVVVOD	IN T PVCT.S	CSGVNGVG	WEGSDS	CSTNAVCC	DENNEGGIIA	TGCT.PA	ST
821	CSTCAT	OCCOST	ESCSSAAVAP		UNT PUCTS	CSDTTCVC	WEGENS	CSANTW CC	DDNSEGGLIS	CCT.PA	ST.
830	CSTCAT	OCCOSV	EPAGSEAVAP		NT PVCVT	CTA TIVV	Vesess	CSENTVCC	DDNSEGGIVS	TGCTPA	57
812	CSTCA	OCCNSV	OSADSLGTAP	TA A TOVITOD	IN T PICT	CSPTSVVC	VeseNA	CSSNAVCC	DNNSEGGTTS	TGCLPA	TT
442	CSTOP	OCCNSV	OKADSPAAAG	TIKSTOVVOD	VDA LVGLT	CSPISVVC	Vesesm	CSGTTVCC	ENNSEGST.TS	TGCVPT	71.
828	CSTGB	KCCNSV	OOANSPSAAA	LIGLIGVVVOC	VDVLVGLT	CDPISVVG	A H-HNS	CSANTVCC	ENNSYGGLTS	IGCV PV	TI
	0010		N 16 The Mar West Street			0 2 2 2 0 1 1 0					

Supplementary Figure 6D. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 2** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150 15
Consensus	KPIAD PYR	XXXXXXZZ	xxxxxx	*****	*****	xxxxxx	X CX TCXX 0	CC XXXXXXXX	XXXXXJ G	TIGX VXX	X X X X I X GJ X	CSPIXXXGX	XXXX CXXXX 🔽	CXNNXXX-	XXGLIXXGO	CX DX X DX MRK
Identity																
011						DNND	ONTRESTO		TO MONTH	ET CUUT S			CC C A OTA C D W	C TO NAME	TOOT TOTO	CHERT THE
877	TAD DVD	CT SNA DR	PORT	AMDONTOFT	TENEVER NDCDCKAVEK	RNNP======	T STOW TOOLS		DV N T N			CSPLAILGVG			TOGET TOTO	
818	KFIADFIK	FUNCOR	APCT. SP	MDDKDMVHD	DAGEAKDATD	SCT PPSCC	- SCN NUCEDIO	CONSUCK NOT	CCV-TCA		PERMIT	CS PI SAUCVO				
776		TPUNNER	MKRGL PP	AAPKRRFSA	HVOOAKRTAP	SAT PPPNGGG	DMSGS NJCP 70	CNOVOOGKS	VT SHIG		DITEAT	CSPTDVICIG	-GN SOSAH PY	CENNS		CTPTT
817		I I I N MON	anno bi i	mir nini on	n' vyynni in r	TPTPOG	- SC S N JCP IT	CEDVR SODD	AS-VATING		D.VTAO.VCIDO	CSPTNUTCAA	SCN ACSAT P	CENNN		CW DV T L
971					TP	VELAARO	T CSTCSA	CESVOSTES	GGLV-ATING	LT GVVV e	DITGOVER	CSPTTVVGVG	SSC SONSNV	SINA	LCCLUSTE	CLPTSI
970				SPWGAP	VTVTVTAPGP	TETAA-S	VCSTCSA	GESVOSATS	EGIV-AT		DITGOVELT	CSPITVIGVG	SGG ECN SN V V	CSDNA	T.C.CT.TSTC	CIPTTI
780						TEPAG	0 CATCNIO	NSVON -S	DPAV-AT	SIGTANG	STITCLYCLT	CSPISVVGVG	SGD ONES PV	T DNS	YG STITSTO	CVPVTL
935			TPNPGEK	RWAVPTT	TATPP	APTNAG	0 CNTCP70	NEW SVOSSS	LSPSIITO G	THE TO TO	DIDGLVCOD	CSPISIIGVG	SGNACSTO PV		VGGTTSTC	CVPVSL
923			TPNPVVK	RWEVPTTKPA	VTOTITVTAP	ASTVT SAG	OONTCSLO	OSTOSATS	AGGEA	MICTANS	GMTGLICSN	CSPITAVGLG	GG SACSAR PV	SNNA	EGETTSTE	CVPIEL
907			TPWG	VTT	VTVTVTAPAP	TVTTVS	OCNTESTS	CCDOVLDSS	EEA NL		GVIGLICID	CSPISVIGVG	SG S ECSAS PV	CENNS	VGGLINIGO	CVPIPL
900			TPWG	APP	VTVTVTAPAP	TVTTVS	OCNTESIO	CCNOVLDSS	EEAGL		GITGLIGIG	CSPISIIGIG	SG S ECSAS PV	CENNS	IGGLINIGO	CIPITL
926			TPWG	ITTITEPGK	VTVTATAPAT	TVTASSG	NCNTGPIO	CONSTMD SKS	SOA NL	LLGIVLG	PREGLICE	CSPISVIGVG	NGG ACSAT PV	T NNN	VGGIIISIG	CIPTIC
824					P	LLAAA-Q	SCNTGPIQ	CON SV EK SD S	AAGSA S	LLGVVLe	DITGSIGIG	CSPLSVVGLG	Q-SSCSASPV	CONNN	VGGLISIG	CLPIEL
451					P	ILAVA-G	SCNTGDV0	CCN SL ED SK S	AAAAG		DVTGLICLO	CSPLTVIGVG	AGSSCS00 PV	CONNN	VGGLVSIG	CVPVOL
641					TP	VARDG-G	NCSTGSVQ	CONSVHDSKE	TAVST		GITGSVGLG	CSPLSVIGVG	SGSSCSSDVV	CONNN	VGGLLSVG	CIPVEL
453					TP	VELAARQ	TCSTCSEO	CONSV VD SN S	ASGNL	LLGIVLG	DITGLICEN	CSPLSVIGAG	SGNACSSNAV	CTNNN	VGGLISIG	CIPIOL
444					TP	VELAARQ	TCSTCSE0	CONSV VD SN S	ASGNLCC	LLGIVLC	DITGLICIN	CSPLSVIGVG	SGNACSSNAV	CTNNN	VGGLISIGO	CIPIOL
446					TP	VELAARQ	TCSTCSE0	CONSVVD SNS	ASG NL III G	LLGIVLG	DITGLICEN	CSPLSVIGVG	SGNACSSNAV	CTNNN	VGGLISIG	CVPIOL
612					TP	LEVRQNG	SCDTGSLQ	CCINTISSIS	AQG NV III G	LLGIVLS	DVIGLICI	CSPISVVGVG	SGN ACTAS PV	CSNNS	VGGLINIGO	CIPIFL
611					тP	LEVRQNG	SCDTGSLQ	CCINTISSIS	AQGNI DIG	LLGIVLS	DVTGLEGLG	CSPISVVGVG	SGN ACTAS PV	CSNNS	VGGLINIGO	CIPITL
610					TP	LEVRQNG	SCDTGSLQ	CCTETISGTS	AQGNVIIIG	LLGIVLS	DVTGLLGLG	CSPISVVGVG	SGN ACTAS PV	CSNNS	VGGLINIGO	CIPIFL
616					TP	LEVRQNG	SCDTCSLQ	CCQSTVSTTS	TNGQDDLCTL	EGLLGVVLG	DVTGL LG LG	CSPISVVGVG	SGN ACTAS PV	CSNNA	VGGLINIGO	CIPIFL
609					ТР	LEVRQSG	SCDTGSLQ	CCD ST IT ST S	ASSNILLC	LLGVVLe	DLTGLVGLN	CSPLSIIG IG	SGN ACTAS PV	CSNTQ	VGGLIGVGO	CIPIFL
436					TP	LEVRQS	SCSSGSVQ	CCDSTISSS	AQAGL 🖪 AG	LLGIVLe	DLTGLIGLI	CSPINVIGIG	SGNACSSNTV	CSNTQ	LGGLIGI	CIPIFL
615					TP	LEVRQNG	SCDTGSLQ	CCESTVEPSD	VPSDPIVGA II	LGLLGIVVe	DLTGLVGLT	CSPISVVG IG	SGD ACTAS PV	CTNSN	VGGLISIG	CVPVSL
910					TP	LEARQDTS	QCNTGSIQ	CCSQTESAS	ESASFLLG	LLGIVL	D ITAL IGL D	CSPISVIGVG	SGSACTASPV	CSNTA	VCGLIGIG	CVPISI
860					SP	TQL SARQ	SCDTGSIQ	CCD STTTT SD	PVTSLLC	LLGIVVS	GVDIP <mark>IGL</mark> N	CSPISVVGVG	-GN SCSAS PV	CTNTDNAF	GL-LSGLIGIG	CVPVNLSL
539 (HGFI)					TP	VR R Q	QCTTGQLQ	CCEST STAND	PAT SELLG	LIGVVIS	DVDALVGL	CSPISVIGVG	SGSACTAN PV	CDSSP	IGGLVSIG	CVPVNV
809					тP	LDTRAAG	KCSTGSLQ	CCN SV QK MD	PATSL	DIGIDVQ	SIDALVGIT	CSPITVVGV	TGN SCNAQAV	CKNNN	VGGAVSLG	CVPVIL
806					TP	LDTRAAG	KCSTGPLQ	CCESTQKATD	PATSVLLA	LIGVNAQ	SVDALVGLD	CSPISVVGV	TGNSCNAQTV	CKNNN	LCGLVSLC	CVPVSLSN
1071						LPTEPAS	QCNTGPVQ	CCNSV SSSSD	PVTSLVLA	LIGVVV	G INV PIGLI	CN PITVIGAG	A-NSCTAQTV	CSNNSFN-	GLIAVGO	CIPIN INL
1008					TPAVM	RRTEPA S	SCSTGPVN	CCNSSGTAKD	GNIAKELA	LIGIVV	D INAL IGV	CSPITVIGAG	G-ASCSSQTL	CEDNKYN-	GIVALGO	CIPVDISL
664				VPNO	GGGGG STPPG	TPGPGPG	QCNTGPIQ	CCESVQRADS	VAAST	LLGVVV2	DISIPIGI	CSPISVIGL	G-NSCSSOPVO	CEDNSFK-	GVIAIGO	CPVNINV
663					G	PLEARQG	QCNTGPVQ	CCNSISTAKD	PATSLLC	LIGIVV	DLNIP <mark>IGL</mark> I	CSPISVIGL	G-NSCSTQPV	CEDNSFN-	GVVAIGO	CIPIN INV
1065						IPA S	SCTTGPIQ	CCNSKELADG	PTASV	LLGVVV	SLOVLVGII	CSPIDVIGLG	S-AACSAE PL	CT NNN FGF	LALSDGITAVG	CVPVDLDL
1059						IPA S	QCISGPVQ	CCNLV ERADG	PTAAA	LIGIVV	DSSILVGI	CFPITIGVG	S-ASCTAOPVO	CODNSH	NGIVAVGO	CVPVDLSL
1041					TPTR	RGGLPAS	SCNTGPIQ	CONTVTTAYD	PTAAE	LIGIVI	N PN ILVGL	CN PITVIG IG	G-SECSSIV	CENNSFG -	GVVSIGO	CVPVDLNL
1063				TCTI	TTVTVTPPAS	TATQ PA S	QCNTGPIQ	den svQs <u>a</u> ss	TPASNVLC		DENATAGE	CSPVTILG GI	F-GG GSAT PV	CENNSFG-	GVIAIGO	CVPVNLSLRK
550					TPLDT	ALPRAAD	QONVSNQ	CONSVQQASS	GPAAL	LLGVVLO	DANATAGED	CSPITVIGEG	N - G G CN S PV	ENNSFG-	SLISIGO	CVPISI
570			G C	VPTTTTKTV	STVSTATIT	ATPIPA S	QCSTAN MQ	HENAL ERADS	TPVGV	LLGVVLO	DEEALIGIT	CSPISIVGVG	QGAQCANQA	ONNT FN -	GLIAIGO	CVPTITQL
202			T PLG S	NPTTTTQTV	TTVTG SGATT	TVTEPA S	SCSSINLQ	OCD SLQSSR S	DPVSK	LIGVVLD	DVTALVGIT	CAPINIVGL	N-A SCNEQAV	CONNT FN -	GITAIGO	CVPINISL

Supplementary Figure 6E. Sequence alignment and corresponding consensus sequence of proteins contained in Region 3 (full sequence)

	1	10	20	30	40	50	60	70	80	90	99
Consensus	CXTGXX	OCCXXXXX -	xxxxxxxxx	(JIX	Gx 🗸 x x x x x x 1	XGJXCSPIXX	XXGXGXXXXCX	XXXXVCCX MNX	XX GL	X X GCX PX)	X L X ERK
Identity			_						_		
											_
907	CNTGSI	SCCDQVLD-		LG	GIVLDGVTGI	LGLDCSPIS	V IGVG SG SEC	SASPVCCENNS	VGGL	INIGCVPI	C L
900	CNTGSI	OCCNQVLD-		G	GIVLEGITGI	LGLGCSPIS	I IG IG SG SEC	SASPVCCENNS-	IG <u>en p</u>	NIGCIPI	C 16
910	CNTGSI	OCC SQTES -	ASSESASI		GIVLEDITAI	IGLDCSPIS	V IGVG SG SAC	ASPVCCSNTA-	VG en	GIGCVPI	SI
612	CDTGSL	OCCTNTIS-		7 I <u>G</u> LG I <u>G</u>	GIVL SDV TGI	LGLGCSPIS	VVGVG SGNAC	ASPVCCSNNS-	VG <u>en p</u>	NIGCIPI	C 16
611	CDTGSL	QCCTNTIS-	S T SAQGN 1		GIVL SDV TGI	LGLGCSPIS	VVGVG SGNAC	ASPVCCSNNS	VG <u>en</u>	NIGCIPI	C 16
610	CDTGSL	QCCTETIS-	GT SAQGNV	7 IILG III	GIVL SDVTGI	LGLGCSPIS	VVGVG SGNAC	ASPVCCSNNS	VG en	NIGCIPI	C 11
616	CDTGSL	QCCQSTVS-	TT STNGQI	DULGTLEGU	GVVLGDVTGI	LGLGCSPIS	VVGVG SGNAC	ASPVCCSNNA	VGGL	NIGCIPI	C II
609	CDTGSL	OCCDSTIT-			GVVLGDLTGI	VGLNCSPLS	IIGIG SGNACI	ASPVCCSNTQ	VG en	GVGCIPI	C 11
436	CSSGSV	QCCDSTIS-	S SAQAGI	G	GIVLGDLTGI	IGLTCSPIN	VIGIGSGNAC	SNTVCCSNTQ-	LGGL	GIGCIPI	C II
923	CNTGSL	OCCQSTQS-	AT SAGGEA		GIAL SGVTGI	IGSNCSPIT	AVGLGGGSAC	GAR PVCCSNNA	EG <u>CL</u>	ISLGCVPI	C 16
971	C STG SA	OCCESVQS-	AESGGLVAI		GVVVGDITGO	VGLTCSPIT	VVGVGSSGSC	ISNVVCCSDNA	LG <u>CL</u>	SIGCLPI	SL
970	C STG SA	QCCESVQS-	ATSEGIVAT		DIVAGDITGO	VGLTCSPIT	V IGVG SGGEC	ISNVVCCSDNA	LGGL	SIGCLPI	C II
780	CATGNL	QCCNSVQN-	ASD PAVAT	ILG SI	GIAVGSITGI	VGLTCSPIS	VVGVG SGDAC	IESPVCCTDNS	YGS	ISIGCVPV	C II
615	CDTGSL	Q C C E S T V E P	SDVPSDPLVGP		GIVVEDLTGI	VGLTCSPIS	VVGIGSGDAC	ASPVCCTNSN-	VGGL	ISIGCVPV	SL
453	CSTGSE	QCCNSVVD -	S N SA SG NI	GLG	GIVLGDITGI	IGLNCSPLS	V IGAGSGNAC	SNAVCCT NNN	VG GL	SIGCIPI	21
444	CSTGSE	OCCNSVVD -		GIG	GIVLGDITGI	IGLNCSPLS	V IGVG SGNAC	SNAVCCTNNN	VGGL	SIGCIPI	21
446	CSTGSE	OCCNSVVD -		JULG	GIVLGDITGI	IGLNCSPLS	V IGVG SGNAC	SNAVCCTNNN	VGGL	SIGCVPI	21
926	CNTG PI	QCCNSTMD -	SK SSQANI	GIG	GIVLGPIEGI	LGLGCSPIS	V IGVGNGGAC	SAT PVCCTNNN	VGGL	SIGCIPI	IL
824	CNTGPI	QCCNSVEK-		1 🛛 – – – – L S 📶	GVVLGDITGS	GIGLGCSPLS	VVGLGQ-SSC	BASPVCCONNN	VGGL	SIGCLPI	81
451		QCCNSLED-	SK SAAAAG	GILSVI	GVNVQDVTGI	IGLOCSPLT	V IGVG AG SSC	SQQ PVCCQ NNN	VGGLV	SIGCVPV	21
641	CSTGSV	QCCNSVHD -	S K ETAV S1	CILGUI	GVSLEGITGS	SVGLGCSPLS	VIGVGSGSSC	SDVVCCQNNN	VGGLI	SVGCIPV	EL
935	CN TG PV	Q CCN SV Q S -		ITQLGII	GIGLGDLDGI	VGQDCSPIS	IIGVG SGNAC	STQ PVCCQ NNN	VGGL	IS IGCV PV	SL
1063	CNTG PI	Q CCNSVQS-	A S ST - PA SN		GVTLTDLNVI	VGLTCSPVT	ILGGIFG-GC	SAT PVCCENNS	FG G V 1	IGCV PV	IL SLRK
550	CNVSNQ	0 CC N SV 0 0 -	ASSG-PAAI		GVVLQDVNVI	VGLDCSPIT	VIGGGNG-GCN	IASPVCCENNS	F G SL	ISIGCVPI	SI
911	CNTGSI	QCCESVQS-	ASSE-TVTN		GVVLSDLDVI	VGLECSPLS	IIGVG SG SACI	ASPVCCENNS	IGGL	SIGCVPI	гь
570	CSTANM	QCCNALER -	ADST-PVGV	7 ILGII	GVVLQDLEAI	IGITCSPIS	IVGVGQGAQC	NQAVCCQNNT	FN GL	AIGCVPI	IQL
568	CSSTNL	QCCDSLQS-	SRSD-PVSF		GVVLDDVTAI	VGITCAPIN	IVGL IN - A SCN	IEQA VCC QNNT	FN G I	AIGCVPI	N I SL
877	CT TG SA	OCCESVES-	ASNP-VAKT		GIVLSDLDVI	VGITCSPIN	ILGLGL-GEC	IEK PVCCENND	FGGL	SIGCSPI	SIEL
818	CNTGPV	QCCNSVQK-	AGTGGVIGA		DVVLDP-EAI	VGLOCSPLS	AVGVGG-SSC	STPVCCENNS	QGGL	SIGCIPI	IL
776	CNTGPV	QCCNQVQQ -	GKSVTS	6 ILG	GVVLDDLEAI	VGIDCSPID	VLGLGG-NSC	SAH PVCCENNS	NS-G	SVGCIPI	гь
809	CSTG SL	QCCNSVQK -	A ND PAT SI	L A	GID VQ SIDAI	VGITCSPIT	VVGVATGNSC	I AQAVCCKNNN	VG G A	SLGCVPV	r L
806	CSTG PL	QCCESTQK-	ATD PAT SV	7 L A L	GVNAQ SVDAI	VGLDCSPIS	VVGVATCNSC	IAQT VCC KNNN	LGGLV	SLGCVPV	SL SN
539 (HGFI)	CTTGQL	QCCESTST -	AND PAT SI	5 [L G 1]	GVVISDVDAI	VGLECSPIS	V IGVG SG SACI	AN PVCCDSSP	IGGLV	SIGCVPV	N V
860	CDTGSI	QCCDSTTT-	T SD PVT SI	. [] LG	GIVVSGVDII	PIGLNCSPIS	VVGVG-GNSC	SAS PVCC TNTDNA	KGL-LS GL	GIGCVPV	NL SL
664	CNTGPI	QCCESVQR-	AD SVAA ST	C 🖬 🗛 🖬	GVVVODLSI	IGITCSPIS	VIGLPG-NSC	S S O PVCCEDN S FI	(GV	A IGC PV	NV
663	CNTGPV	QCCNSIST -	AKD PAT SI	. [] LG	GIVVODLNI	PIGLTCSPIS	VIGL PG - NSC	STQ PVCCEDN SFI	1GVV	AIGCPI	NV
1071	CNTGPV	QCCNSVSS-	S SD PVT SI	J 🗸 🖬 A 🗕 – – – – 🖬 🖬	GVVVQGINVI	IGLTCNPIT	VIGAGA-NSCI	AQTVCCSNN SFI	IGL	AVGCPI	NINL
1059	CISGPV	QCCNLVER-	ADGPTAAA	A LA II	GIVVODSSI	VGITCFPIT	IIGVGS-ASC	AO PVCCO DN SHI	1GIV	AVGCVPV	DISL
1008	CSTG PV	N CCNSSGT -	AKDGNIAB	(ELA	GIVVPDINAL	IGVSCSPIT	VIGAGG-ASC	SQTLCCE DNKYN	1GIV	ALGCIPV	JISL
1041	CNTGPI	QCCNTVTT -	AYD PTAA H	ELFK	GIVIFNPNI	VGLECNPIT	VIGIGG-SEC	SSLVCCENNSFO	3GVV	SIGCVPV	DLNL
1065	CTTG PI	QCCNSKEL -	ADG PTA SV	7 L L G 	GVVVQSLDVI	VGIICSPID	VIGLGS-AACS	A E PLCCT NNN FO	KLAL SD GI	AVGCVPV	DLDL
817	CNTG PI	T CCEDVRS-	ADDA SVAT	LG LI	GVVVQDVTAC	VGIQCSPIN	V IGAA SGNAC	AT PVCCENNNVO	3GL	SVGCVPV	IL

Supplementary Figure 6F. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 3** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 1	0 2	0	30	40	50	60	70	80	90	100	110	120	13
Consensus	X X X S X X X X	*****	x x x x	x x x x x x	*****	CxrGxxQCC	xsvzxxxxxx	X X FIGLIG X 7	. x x x x x x 3	XX CX PIX	- 7 7 - X X 🖸 - X X 🖸	x x Q x VCC x NN	(FNGX IX J)	GCK PINJX LK
Identity														
1067		ASLHO	SPHG===-		SPVAE	СИТСРИДСС	NIVDNFENEDV	ANA LGLVGVVI	OD N IP IG	VNCVPIS.	A7 IG IGE - N SC	SO OP VCCEEN	ENGLVAL	SCTP INAN L
642		SPALAT	TP		IPASQ	CNEGPVQCC	QSISSSNSGT	AS OTOP MONT	SG IN IP IC	VICEPIS	OTO ICE - N SC	SANPVCCONN	FNGWAV	SCIEVNISL
1058		LPO	GPP====		sgg q	CNTG PIQCC	NSVQSSSDPAT	SLILGILLGVVI	QG IN MP IC	LTCNPIT	7 ICAGE - NSC	TAQTVCCENN	FNGVVAVO	GCTP IN INL
631		LPI	EP		AGE	CNEGPVQCC	NSVQSSSNPAT	SL OLCOOLOVV	QG MN MP IC	LICNEIS	7 IC ICC - N SC	TSOTVCCKDNS	FROWATO	SCTPINVSL
650		TI	PAP		QTSQ	CNTGAIQCC	OSVOSSSSSGV	SSILCLEVVI	GG YA VP IC	VICSPIT	7 TGAGE - TSC	DASP VCCENN S	FNGLIA 10	SCTPINLNL
647	-	TI	PAP		QTSQ	CNEGDEQCC	Q SVQ S S SDS SV	ST OLCIOLOVVI	GGHAMP		TOTOS TSC	TAOPVCCENNS	S B N G V I A I O	SCTPINEND
649	PTISERRGO	3 TTT ITVTA	APTA		TTVSQ	CNTG PIQCC	QSTOSSSDPVT	SL OLCIOLC IVI	EGHDVPVC	VICSPIT	TOAGE TSC	ISDP <mark>VCCENN</mark>	PNGVIAIO	SCSPINLNL
648	PTICERRGO	GGATTITVT	APTA		TTVSQ	CNTGPIQCC	OSTOSSBOPVT	SL ALCLES IVI	EGHDMP7G	VICSPIT	7 IGAGE - TSC	ISDP VCC ENNS	FNGVIAIO	GCSPINLNL
1064	ISKPETN	AERLARGLPE	PLPPTRRS-		SGTVGQ	CNTG PIQCC	NSLCKSTDGPI	AADO KIING VV	OD ALVI	LICSPIN	7 TOGOG - NSC	SSQPVCCSNNS	S BEGINNER HO	SCTPININL
567	LPSRETN	AERLARGLPE	PMPPTVRSI	P = = T R VA R NI	KPSGTPSQ	CSTGSLQCC	NSVQSSTSGVV	SA MICHINE P	GGAAMPIC	LTCS PLS	IICIGE - NSC	SOOTVCCBNN	BINGLIN IC	SCTP IN ISL
583	ASHAETN	AVRFARGLPI	PLPPHRRA	TPVAMARRS:	SPSGSPSGGQ	СИТСРИСС	NSVITANHODT	SINCLEVI	-DPSIL76	LOCSPEN	VICVOS - NSC	SO OP VCCENN	INGLIVIC	GCSPIPL
472		IPR				FSTGSLQCC	SSVOKATOPLA	SLIDELLG IVI	GP DLL 70	VICSPIT	VIGVOS-TSC	TOOTVCCTON	BINGLIN IC	SCSPINISL
469		IPR				CSTG3LQCC	SSVOKATOPLA	SLL IGLEG IVI	GP DLL 7G	VICSPIT	VICVOS-TSC	TOOTVCCTON	ENGLIA IO	GCSPINISL
460		IPR				CSTGBLQCC	SSVOKATOPLA	SLL IGLEG IVI	GP DLL 76	VICSPIT	7 IGVGS-TSC	TODTVCCTON	SENGLIA IO	GCSPINISL
475		IPR				CSTGBLQCC	SSVOKATOPLA	SLIDELLG IVI	GP DLL 76	VICSPIS	VIEVES-TSC	TODITVCCTON	BINGLIN IC	GCSPINISL
1056					VLAAASK	CSTGPVQCC	NSVQAANSPTA	VETTE ITE IV	os 7rco7c	VICSPIS	7 LCLGS - NSC	NAOT VCCENN;	ENGVALO	SCTPINVSL
1034	-				VLAAASE	CSTGPVQCC	NSVQAANSPTA	VCTERTEAV	os 7 r co 7 c	ITCNPIS	7 LCLCS - NSC	SAOT VCCENN	S FR G WYA LO	GCTP IN INL
629	S 9				VPGGSGS	СИДСРОСС	NSVQAANSPAA	VCT ICT ICAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	OD ALVO IC	ITCSPIT	-AICLGS-NSC	SSOT VCCENN ;	FNGIIAVO	GCTPINISL
628	S 9				VPGGSGS	СИТСРОСС	NSVQAANSPAA	VET LET LEAAAA	OD ALVO IC	ITCSPIT	-AICLGS-NSC	SSOT VCCENN	FNGIIAVO	GCTP IN ISL
635	S 9				VPGGSGS	СИДСРОСС	NSVQAANSPAA	VCT ICT ICAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	OD ALVO IC	ITCSPIT	-AICLGS-NSC	RSOTVCCENN	SENGIIA VO	GCTPINISL
1054	S 9				VPGGSGS	СИТСРОСС	NSVQAANSPAA	AGLEGIEGVVV	OD ALVOAC	LTCSPIT	VLGLGG-NSC	SAQTVCCENN	FNGIVALO	GCTPINISL
798	TPTP				VPQG-GS	CNTG AIQCC	NSVQSADSPAV	VCTECTEC IAA	OD ALVOAC	VICEPIS	-7 IGVGS-NSC	NAOPVCCENN;	FNGWAL	GCTPINLNL
915	SPAPNGAPI	PPTTTTHPPV1	TTVT	VTAPA	TTTT IPAS - Q	CNEGDAQCC	NSVQSATAPAV	TSLEGLEG IVI	EDINVLVG	LSCOPIS	- 7 IGVGEGANC	VODPVCCENN	FNGLINIC	GCTPINIEL
540	SPAPNGAPI	PPTTTTHPPV1	TTVT	VTAPA	TTTT IPAS - Q	CNEGDAQCC	NSVQSATAPAV	TSLEELE IVI	EDINVLVG	LSCOPIS	7 IGVGGGANC	VO DP VCCENN	FNGLINIC	GCTP IN IFL
485	LPGHPO	S = = = T T T P PV T	TTVT	VTTPP:	STTTIAAGGT	CTTGSLSCC	NOVOSASSPV	TALLGLIG IVI	SDINVLVG	ISCSPLT	VICVGS-SGC	SAQTVCCENTO	PNGLIN I	GCTPINI.
484	LPGHPO	S T T T P PV T	TTVT	VTTPP:	STTTIAAGGT	CTTGSLSCC	NOVOSASSPV	TALLGLEG IVI	SDINVLVG	ISCORT	-7IGVGE-SGC	SAQTVCCENTO	SENGLIN IC	GCTPINI.
483	LPGHPO	S T T T P PV T	TTVT	VTTPP:	STTTIAAGGT	CTTGSLSCC	NOVQSASSPV	TALLGLIG IVI	SDINVLVG	ISCSPLT	VICVCC-SCC	SAQTVCCENTO	FNGLINIC	GCTPINI
480 (SC3)	LPGHPO	S T T T P PV T	TTVT	VTTPP:	STTTIAAGGT	CTEGSLSCC	NOVQSASSPV	TALLGLLG IVI	SDINVLVG	ISCSPLT	-7IGVGE-SGC	SAQT VCCEN TO	PRGLINI(GCTPINI
885	MPGGWC	3GGS 1	TTVT	VTA	ΡΑΤΤΥΤΑΑ SQ	CNTGDIQCC	NSTSSASDPAT	SLLLGLLG IV	QG 7 D V L 7G	IDCSPIS	-715- <u>66</u> -55C	TAEP <mark>VCC</mark> ENNY	FNGLISLO	GCSPINLELK
711	TPWATTO	C P PV 1	TTVT	ITATA	PGTTITSVSQ	CNEGDLQCC	NSVESASSAGA	SLLELGING IVI	SDVIGLVG	ITCSPLS	7 IG IGS-SSC	TAEP <mark>VCC</mark> ENNS	SENGLISLO	GCSPIPL
1060	T P	VAÇ	TGVVAR		DSIPASQ	CSTOSLOCC	NALERADDTAV	GALFELEAAI	QN 7EAL 7G	ITCSPIN	- TLC ICD GGKC	A A Q P VC C Q N N	FNGIIAI	GCVPINISL
697	т 9	vaq	TGVVAR		DSIPASQ	CSTBSLQCC	DALERADSTAV	GVLLGLLGVV	ION VEA IVG	ITCSPIN	ILCICOGATC	AAQPVCCQDN	FNGIIAI	GC7PINISL
1038	TRTSTVTV	DTCTTKTVT	TAPACT		VTQPASQ	CNTG B IQCC	NSVESAKSDPV	TKLLGLLNVVI	SD N IQ76	LTCSPIS	ILCC-GAGC	VAQPVCCQNN	SENGVIA 10	COPVNINL
541	т Р	A1	LKPVAR		DTIPASQ	СИТС І ОСС	NTVENADSPSA	AALEGLEGVV	QG LDVLVG	LTCEPIT	JUCCEGANC	VQDPVCCENN	FNGLINI	GCTPVNLSL

Supplementary Figure 6G. Sequence alignment and corresponding consensus sequence of proteins contained in Region 4 (full sequence)

	1	10	20	30	40	50	60	70	80	88
Consensus	CX-	CEXXOCCXXXXXX	xxxxxx	XJXX	XXXXXXX	CXPIX-VXEXE	- EXSCXXOX	7CCXXNXFNGX	XXXGCXPIN	XXIII
Identity										_
,	-									
1068		IGDIQ CENTV DNS	$H \mathbf{N}KD = I$	KULGULNUG	QDII NV P ICII N	CNPIS-VIGIG	GNECTOOP	CCDDNNFNGI	VAVGCTPVNV	VNL
698	CN-1	IGDIQ CC NTVDKA	SNKD	K ULGUU NUG V	QDII NV P IGI N	CNPIS-VIGVG	-GNSCTQQP	CCEKNNFNGI	VAIGCTPIN	V N
700	CN-1	IGDIQCCDTVDNA	H−− N KD V SI	DALGUL	QDVNVP IGL N	CDSIS-VIGVG	-GNSCTQQP	CCE-NKFNSI	AAIGCTPVN	VNL
1046	CN-1	IGDIQ CCN TVDNA	H – – NKD 🖪 A I	KÖTCITCIA	QDINIP ICI N	CNPIS-VIGIG	-GNSCIQOP	/ CC BNNNF NGV	VALGC TPINV	VNL
1040	CN-1	IGDIQCCNTVDNA	HNKDAI	KQLGILGIVI	QDINIP IGL N	CNPIS-VIGIG	-GNSCTQQP	CCDNNNFNGV/	VALGCTPINV	VNL
1049	CN-1	IGDIQCCOTVDNA	HNKDSI	HALGILDIVV	QDENIP IGE N	CNPIS-VIGIG:	-GNSCTQQP	/CCDBNNFNGI	VALGCTPINV	VNL
1048	CN-	IG DVQCCNTVDNA	HNKDA]	LIASILGIVV	QDANVP IGI Q	CNPID-VIGIS.	-GNSCTQQP	/CCDENNFNGI	VAVGC TPVNV	VNV
1053	CN-1	IGPVQCCNTVDNI	QNEDVA1	NA LGILINI IV	SDVNIP IGL N	CDPISAVVGIG	-GNSCSQQP	/CCBNNNFNGV	VALGCTPINI	LNL
1050	CN-1	IGPIQCCNTVNNI	QNEDVT1	NA LG LL N LVV	QDII NV P IGI N	CDPISAVVGIG	-GNSCSQQP	/CCBNNNFNGI	VAIGC TPINI	LNL
1047	CN-	IGPMQ CCNIV DNA	QNHDVS1	NALGILINIVV	QDIINV P IGIIN	CDPIS-VVGLG	-GNACSOOP	/CCBNNNFNGI	VAIGC TPINI	LNV
1055	CS-	IG PV HCCNSV QAA	NSPTAA(GILGILGIA	QSVTGQ VG VT	CSPIS-ALGLG	GNSCSAQT	/CCBNNNFNG1	VALGCTPINV	VSL
1051	CS-	IGPVQ CC KSLODA	KSPSVA(GLLTGVGVAV	QDVTGQ VGV T	CNPIT-VIGAG	GNSCTSOT	/CCBNNNFNG1	VALGCTPINY	VNL
1061	CN-	TESVQ CC NDIKKA	DDPAVAC	GILGIVGAVV	QDVTAQVGVT	CNPIT-VLGEG.	ANSCSAOT	/CCKDNSFNGV	VAIGCTPINV	VNL
589	CN-	IGSLQCCNSLOSA	NSASIAC	GLLGLLGVVV	GTITGQVGVT	CSPIT-GVGVS	GTSCSEOP	/CCTNNAFNGV	IALGC SPIN	INL
587	CN-	TG SLOCCNSTG SA	TDPA A	KIFAILCINV	EDVTALVGVT	CSPIT-VVGAS	-GSSCSEOP	/CCTNDSFNG1	VALGCAPINI	LNL
999	CS-	IG DLQ CCKSV QSA	DSKSTT:	SIFGILGIVV	GDITGLVGVT	CSPVT-VVGAG	- CAQCNA OA	/CCNDNSFNGI	IALGC TPINI	LNL
575	CSA	TG TAOCCESTOSP	SDLSPSVV	LLGLLGVVV	GDETANVGLT	CSPIT-VIGVG	GTOCNNOV	CCDDNNFNGI	IALGC TPIN	IGL
557	CTA	TG TAQCC DSTQSP	TDLSAPVS	LLGLLGVVV	SQLTGNVGVS	CSPIT-VIGVG	GTQCNNOV	CCNDNNFNGI	VSLGC TPLN	IGL
555	CSA	TG TVOCCASTOSP	SNLGPVIT	GLLGLLGVVI	SDITASVGVT	CSPIT-VIGTG.	GTSCTSOT	CCDNNTFNGV	VALGCTPIN	INL
918	CN-	TGSIOCCNALOSA	GAPG S	TLGILGIK	GD ∛ NA I VG FG	CAPIT-VGGAG	GASCAAOP	CCTGNSFNGI	INIGCTPIN	I
1066	CN-	IGPIQCCNSVOAA	DSTTVT	TLAGILGIVI	SDVTALVGLN	CNPIS-VIGVG	GNSCTOOP	CCTGNNFNGI	IVVGCTPINV	VNL
1000 (VMH2)	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	/CCTGNNFDGI	IVAGCSPIN	IGL
477	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFDGI	IVAGCSPIN	IGL
471	CS-	IGSLOCCSSVOKA	SDPLVG	INVANGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFNGI	IVAGCSPIN	IGL
459	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFNGI	IVAGCSPIN	IGL
476	CS-	IGSLOCCSSVOKA	TDPLASI	ITIGTIGTVI	GPIDLLVGVT	CSPIT-VIGVG	GTSCTROT	/CCTGNSFNGI	IAIGCSPIN	ISL
468	CS-	IGTIOCCNSVOAA	SNPIVGI		GPITGOVGLT	CSPIT-VIGVG	GNSCSAOT	/CCTGNSFNGI	LVVGCSPIN	INL
458	CT-	IGSLOCCNSVOAA	SNPVVGI	LIAGINIGIVI	GPITGOVGLT	CSPIT-VIGVG	GTSCSAOT	CCNGNSFNGI	IVVGCSPVN	ISL
586	CN-	IGTLOCCNSVOOA	SNPVVGI	LIAGINGIVI	GPHTGOVGLT	CSPIS-VIGVS	GNSCSAOP	CCTGNSFNGI	LVVGCTPINI	LNL
920	CN-	IGTLOCCNSVOSS	SDPVTSI	LILGII	GGIDIPIGIO	CTPIT-VIGVG	GANCVOOP	/CCTGNTFNGI	VIVGCSPINI	LGL
921	CN-	IGPLOCCNSVOOA	DAPGAA	OLLKTIGVVV	OGTTTMVGIN	CSPIP-ILGAA	GTKCTOOP	CCBNNNYNGI	VNIGCSPIN	GDL
710	CN-	IGDLQCCNSVESA	SSSGVA	LGIGGIVI	GDITGLVGIT	CSPLS-IIGLG		CCENNTFNGI	ISLGCSPIT	L
552	CT-	GELNCCOOVEST	ASNPVO	ALSCILCVAT	GALGLNVGLG	CSPIS-VIGIG	GNSCSAOT	CCBGNTFNGI	IALGCMPIN	IGL
643	CN-	TG SLOCCOTLSNS	SNSGMT	TISGILGIV	P-ANV PVGUT	CNPISILVGIG	GNSCSAOP	CCOGNNENGI	IVLGC TPVNI	LNL
916	CD - Z	GPVOCCNSVOKS	SNPSVAI	KILSGLAVPI	OGISVPICIT	CSPLN-LLALG	GNSCASOP	CCENNNFNGI	VAIGCTPID	GAT

Supplementary Figure 6H. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 4** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30	40	50	60	70	80	90	100	110 116
Consensus	SASAHXXXXXX	************	XXXXXX CX CXXXXX	X OCCXXXXXXX -	-xxxxxxx	XXIICXXXXXXXX	X X X X X X X I	X - 7X CXC	GX SCX X OX V	CCXXXXXENGXX	XXGCXPINXXEL
Identity		_									
								_			
1068		V P T P G G H C	GHGGS <mark>KCNTG</mark> D	IQCCNTVDNSH-	-NKDEK	LGLLNIGVQDLN	7 P IGL N C N P I	S-VIGIG-	GNSCTQQPV	CCDDNNFNGLV	AVGCTPVNVNL
698		VAT	GSS <mark>KCNTG</mark> D	IQCCNTVDKAS-	-NKDIQKI	LGLLNIGVQDLN	7 P IGL NCN PI	S-VIGVG-	GNSCIQQPV	CCEKNNENGLV	AIGCTPINVN
1046		APGG	- HG D A K C N T G D	IQCCNTVDNAH-	-NKD AKQ	LGLLGLVIQDUNI	IPIGLNCNPI	S-VIGIG	GNSCIQQPV	CCBNNNFNGVV	ALGCTPINVNL
1040		APGG	- HG D A K C N T G D	IQCCNTVDNAH-	-NKD AKQ	LGLLGLVIQDUNI	IPIGLNCNPI	S-VIGIG-	GNSCTQQPV	CCDNNNFNGVV	ALGCTPINVNL
1049		TPGG	- HD G A K C N T G D	IQCCQTVDNAH-	-NKD SHA	LGLLDIVVQDIN:	IPIGLNCNPI	S-VIGIG-	GNSCIQQPV	COBNNENGLV	ALGCTPINVNL
700		APGG	- HG D A K C N T G D	IQCCDTVDNAH-	-NKDZSDA	LGLUNIIVQDVN	7 P IGLNCDS	S-VIGVG-	GNSCTQQPV	CCEN-KENSLA	AIGCTPVNVNL
1048		PTYDPYGHHI	DTPVAKCNTGD	VQCCNTVDNAH-	-NKDIAL	A S FLGIV V QD A N	7 P IGIQCNPI	D-VIGIS-	GNSCIQOPV	CCDENNFNGLV	AVGCTPVNVNV
1053		APHGSPHG	- SPVAQCNTGP	VQCCNTVDNIQ-	-NEDZANA	LGLLNIIVSDVNI	IPIGLNCDPI	SAVVGIG	GNSCSQQPV	CC B N N N FN G VV	ALGCTPINLNL
1050		APHG	- SPVAQCNTGP	IQCCNTVNNIQ-	-NEDVTNA	LGIUNIVQDUN	7 P IGLNCD PI	SAVVGIG	GNSCSQQPV	CCENNNFNGLV	AIGCTPINLNL
1047		APYG	- SPVAQCNTGP	MQCCNIVDNAQ-	-NHD SNA	LGLLNIVQDLN	7 P IGL N C D PI	S-VVGLG	GNACSQQPV	CCBNNNFNGLV	AIGCTPINLNV
643		LPSPAAAT	FIPASOCNTGS	LQCCQTLSNSS-	-NSG7TT	SGILGIVIP-ANV	/ P VGL ECN PI	SLVGIG	GNSCSAOPV	CCQGNNFNGLI	VLGCTPVNLNL
1066		P G G	- QS D S <mark>C N T G</mark> P	IQCCNSVQAAD-	-STTVTT	AGTIGIVISDVT	AL VGL NCNPI	S-VIGVG	GNSCIQOPV	CCTGNNFNGLI	VVGCTPINVNL
1055		VI	LAAASECSTGP	VHCCNSVQAAN -	-SPTAAG	LGLLGIAIQSVT	QVGVTCSPI	S-ALGLG	GNSCSAQTV	CCBNNNFNGIV	ALGCTPINVSL
1051		VI	LAAAHE <mark>CSTG</mark> P	VQCCKSLODAK-	- SPS7AG	LTGVGVAVQDVT(QVGVTCNPI	T-VIGAG-	GNSCTSQTV	CCENNNENGIV	ALGCTPINVNL
1061		(GAQDSECNTES	VQCCNDIKKAD -	- DPAVAG	LGLVGAVVQDVT	AQVGV TCNPI	T-VLGGG-	ANSCSAOTV	CCKDNSFNGVV	AIGCTPINVNL
589		TPTRRNI	SPSNOCNTGS	LQCCNSLQSAN-	-SASTAGT	LGLLGVVVGTIT(QVGVTCSPI	T-GVGVS-	GISCSEQPV	CCTNNAFNGVI	ALGCSPININL
587			SP S NO C NT G S	LQCCNSTGSAT-	-DPALAK	FALCINVEDVTI	LVGVECSPI	T-VVGAS-	GSSCSEOPV	CC IND SENGIV	ALGCAPINLNL
999		TPTRRNH	PAAGS CSTGE	LQCCKSVQSAD-	-SKSCTSC	FGJIGLVVGDIT(LVGVECSPV	T – V V G 🗚 G -	EAQ ON A QAV	CCNDNSFNGLI	ALGCTPINLNL
575		LPRGGO	G Q G A A C S A T G T i	A QCCE STOSPSI	LSPSVT	LGLLGVVVGDLT	AN VGL ECSPI	T-VIGVG-	GT Q CN N QV V	CCDDNNFNGLI	ALGCTPINIGL
557		LPRGGO	GDGAACTATGT	A QCCD STOSPTE	LSAPVST	LGLLGVVVSQLT(NVGVSCSPI	T-VIGVG-	CTQCNNQVV	CCNDNNENGLV	SLGCTPLNIGL
555		LPR1	FDGAACSA <mark>FG</mark> T	VQCCASTOSPSN	LGPVCTGC	LGLLGVVISDLT	SVGVECSPI	T-VIGGG-	GISCISQIV	CCDNNTFNGVV	ALGCTPININL
918	SASAHPAAAA	RAAAKPSATTKPS	SQSAAA CNTGS	IQCCNALQSAG-	- APGISVI	LGLLGIKVGDVNI	A I VGFGCAPI	T – VGGAGA	GASCAAOPV	CCTGNSFNGLI	NIGCTPINI
1000 (VMH2)		IPRTD-	TPS CSTGS	LQCCSSVQKAS-	-DPL V GI	VALLGIVLGPLDI	NVGLECSPI	T-VIGVG-	GISCIQQIV	CCTGNNFDGLI	VAGCSPINIGL
477		IPRTD-	TPS CSTGS	LQCCSSVQKAS-	-DPLVGII	VAJEGIVEGPEDI	INVGLECSPI	T-VIGVG-	GISCIQQIV	CCTGNNFDGLI	VAGCSPINIGL
471		IPRTD-	TPS CSTGS	LQCCSSVQKAS-	-DPL V GI	VALLGIVL GPL DI	IN VGLECSPI	T-VIGVG-	GISCIQQIV	CCTGNNFNGLI	VAGCSPINIGL
459		IPRTD-	TPSCSTGS	LQCCSSVQKAS-	-DPLVGII	VALLGIVLGPLDI	NVGLECSPI	T-VIGVG-	GISCIQQIV	CCTGNNFNGLI	VAGCSPINIGL
476		ILWTD-	TPSCSTGS	LQCCSSVQKAT-	-DPLASL	IGILGIVEGPEDI	LVGVECSPI	T-VIGVG-	GESCIRQEV	CCTGNSFNGLI	AIGCSPINISL
468		IPRTD-	TPSCSTGT	IQCCNSVQAAS-	-NPIVGLL	AGLIGIVIGPIT	QVGLECSPI	T-VIGVG-	GNSCSAQEV	CCTGNSFNGLL	VVGCSPININL
458		IPRTNI	P-PAPTCTTGS	LQCCNSVQAAS-	-NPVVGL	AGLIGIVIGPIT	GOVGLECSPI	T-VIGVG-	GESCSAQEV	CCNGNSFNGLI	VVGCSPVNISL
586		TPVPTDGS	SGPSDQCNTGT	LQCCNSVQQAS-	-NPVVGL	AGLLGIVLGPIT	GOVGLECSPI	S-VIGVS-	GNSCSAQPV	CCTGNSFNGLL	VVGCTPINLNL
710	TSPVT	TVTVTATAPGST	/TTVSQCNTGD	LQCCNSVESAS-	- 🛚 🛛 🖓 🗛 D 🖬	LGILGIVIGDIT(IL VG I C S PI	S-IIGLG	SSSCTABPV	CCENNTFNGLI	SLGCSPIFL
552		GETO	GCTTGB	IN CCQQVESTA-	SN P VQA 🖬	SGLLGVALGALGI	IN VGLGCSPI	S-VIGIG-	GNSCSAQEV	CCEGNTFNGLI	ALGC MPINIGL
921		TPTPSPEML	A A R G G Q C N T G P	LQCCNSVQQAD-	- A P G A A Q 🖥	LKTLGVVVQGTT	MVGINCSPI	P-ILGAA1	GT K CE Q QP V	CC B N N N YN G L V	NIGCSPINGDL
920		TPTPSSESLI	I AR D GOCNTGT	LQCCNSVQSSS-	-DPVTSL	LGLLGVVLGGID	PIGIQCTPI	T-VIGVG	GANCVQQPV	CCTGNTFNGLV	TVGCSPINLGL
916		T PV 1	RWKGOCDAGP	VQCCNSVQKSS-	-NPS AKI	LSGIAVPLQGIS	PIGLTCSPI	N-LLALG	GNSCASOPV	CC B N N N FN G L V	AIGCTPIDLGAL

Supplementary Figure 61. Sequence alignment and corresponding consensus sequence of proteins contained in Region 5 (full sequence)

	1	10	20	30	40	50	60	70	80	88
Consensus	CX-	CCXXXXCC XXXXXX	xxxxxx	XJXX	XXXXXXX	CXPIX-VXEXE	- EXSCXXOX	7CCXXNXFNGX	XXXGCXPIN	XXIII
Identity										_
,	-									
1068		IGDIQ CENTV DNS	$H \mathbf{N}KD = I$	KULGULNUG	QDII NV P ICII N	CNPIS-VIGIG	GNECTQOP	CCDDNNFNGI	VAVGCTPVNV	VNL
698	CN-1	IGDIQ CC NTVDKA	SNKD	K ULGUU NUG V	QDII NV P IGI N	CNPIS-VIGVG	-GNSCTQQP	CCEKNNFNGI	VAIGCTPIN	V N
700	CN-1	IGDIQCCDTVDNA	H−− N KD V SI	DALGUL	QDVNVP IGL N	CDSIS-VIGVG	-GNSCTQQP	CCE-NKFNSI	AAIGCTPVN	VNL
1046	CN-1	IGDIQ CCN TVDNA	H – – NKD 🖪 A I	KÖTCITCIA	QDINIP ICI N	CNPIS-VIGIG	-GNSCIQOP	/ CC BNNNF NGV	VALGC TPINV	VNL
1040	CN-1	IGDIQCCNTVDNA	HNKDAI	KQLGILGIVI	QDINIP IGL N	CNPIS-VIGIG	-GNSCTQQP	CCDNNNFNGV/	VALGCTPINV	VNL
1049	CN-1	IGDIQCCOTVDNA	HNKDSI	HALGILDIVV	QDENIP IGE N	CNPIS-VIGIG:	-GNSCTQQP	/CCDBNNFNGI	VALGCTPINV	VNL
1048	CN-	IG DVQCCNTVDNA	HNKDA]	LIASILGIVV	QDANVP IGI Q	CNPID-VIGIS.	-GNSCTQQP	/CCDENNFNGI	VAVGC TPVNV	VNV
1053	CN-1	IGPVQCCNTVDNI	QNEDVA1	NA LGILINI IV	SDVNIP IGL N	CDPISAVVGIG	-GNSCSQQP	/CCBNNNFNGV	VALGCTPINI	LNL
1050	CN-1	IGPIQCCNTVNNI	QNEDVT1	NA LG LL N LVV	QDII NV P IGI N	CDPISAVVGIG	-GNSCSQQP	/CCBNNNFNGI	VAIGC TPINI	LNL
1047	CN-	IGPMQ CCNIV DNA	QNHDVS1	NALGILINIVV	QDIINV P IGIN	CDPIS-VVGLG	-GNACSOOP	/CCBNNNFNGI	VAIGC TPINI	LNV
1055	CS-	IG PV HCCNSV QAA	NSPTAA(GILGILGIA	QSVTGQ VG VT	CSPIS-ALGLG	GNSCSAQT	/CCBNNNFNG1	VALGCTPINV	VSL
1051	CS-	IGPVQ CC KSLODA	KSPSVA(GLLTGVGVAV	QDVTGQ VGV T	CNPIT-VIGAG	GNSCTSOT	/CCBNNNFNG1	VALGCTPINY	VNL
1061	CN-	TESVQ CC NDIKKA	DDPAVAC	GILGIVGAVV	QDVTAQVGVT	CNPIT-VLGEG.	ANSCSAOT	/CCKDNSFNGV	VAIGCTPINV	VNL
589	CN-	IGSLQCCNSLOSA	NSASIAC	GLLGLLGVVV	GTITGQVGVT	CSPIT-GVGVS	GTSCSEOP	/CCTNNAFNGV	IALGC SPIN	INL
587	CN-	TG SLOCCNSTG SA	TDPA A	KIFAILCINV	EDVTALVGVT	CSPIT-VVGAS	-GSSCSEOP	/CCTNDSFNG1	VALGCAPINI	LNL
999	CS-	IG DLQ CCKSV QSA	DSKSTT:	SIFGILGIVV	GDITGLVGVT	CSPVT-VVGAG	- CAQCNA OA	/CCNDNSFNGI	IALGC TPINI	LNL
575	CSA	TG TAOCCESTOSP	SDLSPSVV	LLGLLGVVV	GDETANVGLT	CSPIT-VIGVG	GTOCNNOV	CCDDNNFNGI	IALGC TPIN	IGL
557	CTA	TG TAQCC DSTQSP	TDLSAPVS	LLGLLGVVV	SQLTGNVGVS	CSPIT-VIGVG	GTQCNNOV	CCNDNNFNGI	VSLGC TPLN	IGL
555	CSA	TG TVOCCASTOSP	SNLGPVIT	GLLGLLGVVI	SDITASVGVT	CSPIT-VIGTG.	GTSCTSOT	CCDNNTFNGV	VALGCTPIN	INL
918	CN-	TGSIOCCNALOSA	GAPG S	TLGILGIK	GD ∛ NA I VG FG	CAPIT-VGGAG	GASCAAOP	CCTGNSFNGI	INIGCTPIN	I
1066	CN-	IGPIQCCNSVOAA	DSTTVT	TLAGILGIVI	SDVTALVGLN	CNPIS-VIGVG	GNSCTOOP	CCTGNNFNGI	IVVGCTPINV	VNL
1000 (VMH2)	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	/CCTGNNFDGI	IVAGCSPIN	IGL
477	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFDGI	IVAGCSPIN	IGL
471	CS-	IGSLOCCSSVOKA	SDPLVG	INVANGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFNGI	IVAGCSPIN	IGL
459	CS-	IGSLOCCSSVOKA	SDPLVG:	INVALGIVI	GPEDLNVGLT	CSPIT-VIGVG	GTSCTOOT	CCTGNNFNGI	IVAGCSPIN	IGL
476	CS-	IGSLOCCSSVOKA	TDPLASI	ITIGTIGTVI	GPIDLLVGVT	CSPIT-VIGVG	GTSCTROT	/CCTGNSFNGI	IAIGCSPIN	ISL
468	CS-	IGTIOCCNSVOAA	SNPIVGI		GPITGOVGLT	CSPIT-VIGVG	GNSCSAOT	/CCTGNSFNGI	LVVGCSPIN	INL
458	CT-	IGSLOCCNSVOAA	SNPVVGI	LIAGINIGIVI	GPITGOVGLT	CSPIT-VIGVG	GTSCSAOT	CCNGNSFNGI	IVVGCSPVN	ISL
586	CN-	IGTLOCCNSVOOA	SNPVVGI	LIAGINGIVI	GPHTGOVGLT	CSPIS-VIGVS	GNSCSAOP	CCTGNSFNGI	LVVGCTPINI	LNL
920	CN-	IGTLOCCNSVOSS	SDPVTSI	LILGII	GGIDIPIGIO	CTPIT-VIGVG	GANCVOOP	/CCTGNTFNGI	VIVGCSPINI	LGL
921	CN-	IGPLOCCNSVOOA	DAPGAA	OLLKTIGVVV	OGTTTMVGIN	CSPIP-ILGAA	GTKCTOOP	CCBNNNYNGI	VNIGCSPIN	GDL
710	CN-	IGDLQCCNSVESA	SSSGVA	LGIGGIVI	GDITGLVGIT	CSPLS-IIGLG		CCENNTFNGI	ISLGCSPIT	L
552	CT-	GELNCCOOVEST	ASNPVO	ALSCILCVAT	GALGLNVGLG	CSPIS-VIGIG	GNSCSAOT	CCBGNTFNGI	IALGCMPIN	IGL
643	CN-	TG SLOCCOTLSNS	SNSGMT	TISGILGIV	P-ANV PVGUT	CNPISILVGIG	GNSCSAOP	CCOGNNENGI	IVLGC TPVNI	LNL
916	CD - Z	GPVOCCNSVOKS	SNPSVAI	KILSGLAVPI	OGISVPICIT	CSPLN-LLALG	GNSCASOP	CCENNNFNGI	VAIGCTPID	GAT

Supplementary Figure 6J Sequence alignment and corresponding consensus sequence of proteins contained in **Region 5** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30	40	50	60	70	80	90	100	110	120	130	140
Consensus	GPT FR PGN XXX XXX	*****	*********		XX 🕏 XX 🗟 XX X	CXXXXXXX	xxxxxx x	*****	XXXX 🗙 XX	JXXXXXX	XX CXXXX 7CC	xxXXX G	XXX X 10 X 2X	ax xx
Identity			_											<u> </u>
990		SPTP	GGYPDSTTV	v	SQ ONV DE 7	ICONTOQT PD	HTNAAAGG	G- АЛА № 7G	-ALEGFDCT	BSVIJIGG	NNCAAOPVCC	EANESIG	-LINALSCOPI	NVNE
489		SPTP	GGYPDSTTV	V	SQ ONV DE 7	ICONTOOT PD	HTN-AAGG	G-AAAN G	-AL DGFD TI	BSVIJIGG	NNCAAOPVCC	EVNESL 6	-LINALSCOPI	NVNE
737		TPAP	G-KPKAS		SQ ODV OF IF	ICCNTQQT PD	HTSAAASG	G- IP IN G	-AF CFDCT	BSALOVGG	NNCAAOPVCC	TSNOST 6	-LINALDCSPV	NVNE
988		VPRT	DPPPTTGSQ		CT AVG GF 7	COTLIGGS	DPVVGL	GLLGIDISL	EGVVGLNCSI	PETVI366-	NSCSS OTA CC	TSNSZNG	IIVIDCSPI	NIDL
738		VPRT	DPPPTTLSQ		CTAVG GF74	CHTLINGS	DPTVRL	GLLDIDESLI	BGIVGLECS	TTVIGG-	NSCIS OFA CC	T 6 N N 7 7 6	-GLLVLGCSPI	NVISI
629			APTSTGGS S-		CNASG GF 7	CONSTESAS	NISGSISC	GILGVNPAT	TGL IGL QCG	ISVVGG	TSCSS PVCC	5 6 N N F N G	LV7LGCSPV	NV OL
987			AP PAP VTN N-		SQ CNV GT 7	ICONSAQD VE	SSAVQT	HSLGVDZGGV	NGLVGLTCN	VLG G	SSCEA DPVCC	TONSFOG	VIVGCSPI	NVSL
974		ADEPAAAP	AP PPP VINN-		SQCDVCK7	D CC D S MOD VR	SSAVQTA	HSLGVDEGG	TGLVGLTCN	VLG G	SSCEADPVCC	DKNHSGG	LINVGCSPV	NVSL
986		APGGEP	AP-STGGQN-		SQ CDV CK II	CCESTODVK	SSAVQN	GLLGIDIGGI	TGIVGVNGG	ILG	SSCSA OPVCC	TSNSFHG	LIVVGCSPI	NVSL
733		APGGEP	AP-STGGQN-		SQ SDV SK I	COSTODAR	SSAVQT	GLLGIDIGGY	TGLVGLTCN	GLSG	SSCIADRVCC	TSNSFDG	LIVVGCSPI	NVSV
729 (SC16)		VP	RDVNGGTPP		KSCSSCP7	CONKTEDSK	HLDKGTTAL	GLLNIKIGD	KDL VGL N SI	STSVISVGG	NSCSADIFCC	TNT YQH 3	LVNVGCTPI	NIGL
728		VP	RDVNGGTPP		KSCSSCP7	CONKTEDSK	HLDKGTTAL	GLLNIKZGDZ	KNLVGLNCSI	PISVIGVGG	NSCSADEVCC	TNTYQH G	LVNVGCTPI	NIGL
659			TPVVARTEN			ICCESTFS SN	HPSVST	GLFGF - 7GNL	GNS IGISCS	NIGSLCC-A	PNCNODE7CC	TSNOZNG	LIAF SCIPF	NFGF
495			TPVVARTEN			ICCESTFS SN	HPSVST	GLFGF - 7GNL	GNS IGISCS	LNIGSLGG-A	PNCNODEVCC	T 6 N 0 Z N 6	LINF SCIPF	NFGF
493			TPVVARTEN			ICCESTFS SN	HPSVST	GLFGF - 7GNL	GNS IGISCS	I I I G SLG G - A	PNCNODEVCC	TSNQZNG	LIAF SCIPF	FGF
770		APPA	AQ PQS AQP Q	NAQPQ	NAQPKCNTCSPF	COSTYTNE	PTKLAS S	KSLDIA PS7	TGL EALT CT	ITGL GT G IG	AV CEQ KEVCC	TDNGZNG	VV LGC SPL	DLDS
763		APPA	AQ PQS AQP Q	NAQSQ	NAQPKCNTCSPF	COSTYTNE	PTKLASIS	KSLD IA PS7	TGL BALT TI	ITGL GT GTC-	AV CEQ REVCC	TONGZNG	VV LGC SPL	DLDS
768		APPA	AQ PQN AQP Q	NAQPONTOPO	NTQ PK CNT CSP F	COSTYTNE	PTKLAS S	KSLGIA PS7	TGL BALT TI	ITGL GT ITC-	AV CTE KPVCC	TONGING	VV LGC SPL	DLDS
758		APGI	LEVRTLADQ		CrtGs 7	CKSTIQST	SMSAAIG	GLLG IVF TP	OGL VGL S TI	TLVRASAG-	TV CTO OPVCC	TSNTYSG	SIN IGC SP I	NFNS
876			AAGGAG		Q ENT CSAY	CNOVESVO	DKSVTKQL	TALCVVVEG	TOONSLDSI	N-VLSTGAG	TCNOHPVCC	TNNK FF 6	LINLGCSPI	NVNL
546			GPLSVRD		OCNT GT I	COVVOUAS	YYQSAFQE	IGEGEL AG	TGO IGT O CSI	S-VVGASNG	AQCNA OF FCC	TNTOFN G	LIN IGC MP I	NVNA
716	IPS PVD	VVR-REGAATS	TSSTSSSSS		SQ ENT CS	CCEST-HSA	SSVETASA	PLSILAAL-G	GGL IGV N SI	APIGVOSC-	ASCAO BRICC	SEDSES	G-LINVGCSPI	NVNP
715	GPTFRPGN TTNHLP	VV ROKPCANNI	TTETTAS		TM CNT GT I	CSST-QSI	SDSMSL	GLLPLGAV-	EGL VGMC SI	LPMGV35G-	TSCNO BRVCC	SEDTFEEPS	G- ILNLGC SP I	NALK
608			TPLPGGSS			COQTYSST	SSEASL	SIVCLDESC	TGS IGS QCS	SAIGLOSG-	SSCIQ	5 N N N Z 0	G-LIVVGCSPI	NL
581			TP IPD AAS S		PO ONT OP I	COSVYOS	TTSHSIZA	NLVGLD70SL	TAS IGT QCS	TVGGLAAG-	AK CSO OPVCC	S S N N S S	G-LIVVGCSPI	INL
661			IPS GAP T-		CATOP IC	CERVYES	TTETSLET	DLLGLNEDGE	LGGEATGES	SVVGIGG-	NKCAHRPVCC	TONKENG	LVNVGCVPV	NVNE
657			NP IPN SEGV.		CNT GP 7	CETRFSAC	SREANL	SLLGLDIGGI	LGD BASG SI	SVVGVGCG-	TRCSSAPVCC	TONKEN G	LINVGCVPV	EV GL
956		I	P		OS SP COTAP 7	CONVAPON	SDHMRE	KALKMN D-P	DTHYOTGAL	ANPVGTOG	TTCSSSPMCC	DGNS	LISLGCTPI	PUDL
670	PAGGPN	AE RLARGL PPL	PP VRR HAT P	AHLAARAQPS	SSS SQ CNC GT I	CONSVASSN	DAVPKL	SILNLGGG-	NTIVGMOCT	NAL GVECC-	SSCIG OF 7CC	S G N D E N G		SIGA

Supplementary Figure 6K. Sequence alignment and corresponding consensus sequence of proteins contained in Region 6 (full sequence)

	1	10	20		30	40		50		60		70		80		90 92
Consensus	СВХСХ	XXCC	XXXXXXXX	XXXXXXX	XXIIXXX	XXXXXX	XXexx	CXXJXXX	GXXXC	XXOXX	XXVCC	XXXXX	XX – XX –	XXXX	K C X P X N	XXXI
Identity							_					_				
			TROOMDD			AN TO A								Edit a C		TEATT
990			TQQTPD	HTNAAAG	GLIGAA	AN -VGA	LIGFD	TPLSV	GI-GG	NNCAA	OPVCC	SANER	TGL	NA	13CSPIN	VNL
489			TQQTPD	HTN-AAG	GLIGAA	AN-VGA	LIGFD	TPLSV	GI-GG	NNGAA	QPVCC	SANER	TGL	NA	15C5PIN	VNL
/3/	GDVGE	TH CC	TQQTPD	HTSAAAS	GLIGIP	IN-GA	FUGFD	TPISAL	GV-GG	NNGAA	QPVCC	GNOR	TGL	INA	DCSPVN	VNL
988	GTAVGEF	VY CC	MTLIGGSD	PVVGLLA	GLUGID	LSLEEG	AAGT N	SPLTVL	GGe	NSOSS	OTACC	GNSY	NG		DCSPIN	DL
738	CTAVGCF	VY CC	HTLIHGSD	PTVRLIA	GLEDID	LSLIEG	IVGLI	SPLTVL	GGe	NSCIS	OTACC	IGNNY	Y G G		GCSPIN	VEL
729 (SC16)	CSSCP	VY CCI	NKTEDSKHLD	KGTTALL	GLUNIK	IGD <u>I</u> KD	LVG LN	SPUSVU	GV-Ge	NSCSA	QTVCC:	I – NTY	QH G	LVN	GCTPIN	IGL
728	CSSCP	A A CC	NKTEDSKHLD	KGTTAL	GLUNIK	V G D U K N	LÅG TN	SPUSV	GV-GC	NSCSA	QUVCC	$\mathbf{I} - \mathbf{N}\mathbf{T}\mathbf{Y}$	QH G	LVN	/GCTPIN	IGL
987	CNV GT	LHCC	N S AQDVKS	SAVQTLA	HS LGV D	LGGVNG:	LVGLI	NV	G@	SSCSA	QPVCC:	IGNSE	G G	VIV	/GCSPIN	VSL
974	CDV GK		Q S MQDVKS	SAVQTLA	HS LGV D	LGGVTG	LVGLI	NVL	G@	SSCSA	QPVCC	DKNHF	G G	LIA	/GCSPVN	VSL
986	CDV GK	DCC	E S TQDVKS	SAVQNLA	GLLGID	LGGITG	IVGVN	CGII	GG	SSCSA	QPVCC	IGNSE	H G	LIV	/GCSPIN	VSL
733	CDVCK		Q S TQDAKS	SAVQTLA	GLLGID	LGGVTG	LVGLI	NGI	SG	SSCTA	QPVCC	IGNSE	D G	LIV	/GCSPIN	VSV
629	CNASGGT		STESASNIS	GSISGLL	GILGVN	PATETG	LIG LQ	G-ISVV	G 🗗	TSCSS	QPVCCS	GNNF	NG	LVVI	LGCSPVN	VGL
659	CNT GS	LHCC	E S TFSSNH	PSVSTLA	GLFGFV	GN – 🖬 GN	SIGIS	SALNIG	SLGGA	PNCNQ	QIIVCC	GNQY	NG	LIAI	GCTPFN	FGF
495	CNT GS	LHCCI	ESTFSSNH	PSVSTLA	GLFGFV	GN – 🖬 GN	SIGIS	SALNIG	SLGGA	PNCNQ	QIIVCCI	IGNOY	NG	LIAI	GCTPFN	FGF
493	CNT GS	LHCC	ESTFSSNH	PSVSTLA	GLFGFV	GN – 🖬 GN	SIGIS	SAUNIG	SLGGA	PNCNQ	QUVCC	GNOY	NG	LIAR	GCTPFN	FGF
670	CNGGT	TKCCN	SVASSND	AVPKLLS	SILNLG	LG – IN T	IVGMO	TNINAL	GVGGG	SSCTG	QUVCCS	GNDE	NG	VITZ	GCTPIS	GA
770	CNTCS	PFCCI	S TYTNDP	TKLASLS	KSIDIA	LPSVTG	LALT	TPITGI	GTGT	AVOTO	KPVCC	DNGY	NG-	VVNI	GCSPLD	DS
763	CNTGS	PFCCI	S TYTNDP	TKLASLS	KSIDIA	LPSVTG	LALT	TPITGI	GTGT	AVOTÕ	KPVCC	DNGY	NG-	VVNI	LGCSPLD	DS
768	ONTGS	PFCC	S TYTNDP	TKLASLS	KSIGIA	LPSVTG	LIALT	TPITGI	GTGT	AVCTE	KPVCC	DNG	NG-	VVNI	GCSPLD	DS
758	CTTGS	TOCC	STIOSTS	MSAAILG	GLUGIV	FTPHOG	LVGLS	TPITLY	RAGA	TVCTO	OPVCC	IGNTY	SG-	SIN	IGCSPIN	FNS
716	GNTGS		STHSA	SSVETAS	APUSTL	AALGGG	LLGVN	SPILAPI	GVGS	ASCAO	EPICCS	GDSF	SG-	LINK	GCSPIN	VNP
715	GNTGT		SSTOSL	SDSMSLL	SGULPL	GAVEEG	L∛GMG	SPILPM	GVGS	TSONO	EPVCCS	GDTF	GGPSG-	TUNI	GCSPIN	ALK
608	ONT es		DOTYSSTS	SEASLIA	STVGLD	LSGVTG	STESO	SPUSAL	GLGS	SSCTO	OPVCCS	NNNY	OG-	LTV	GCSPIN	11
581			SVYOSOT	TSHSTLA	NLVGLD	VOSUTA	STGTO	SPILTVG	GLAAR	AKCSO	OPVCCS	GNNE		TTV	GCS PT N	T.
876	ONT es	AVCON	NOVKSVOD	KSVTKOL	TAILGVV	VEGUTC	OACTO	SPUNVI	STGAR	-TONO	HPWCC	NNKE	TG	TTN	GCSPIN	WNT.
546				VOSAFOE	TGUGET	LACUTC	OTETO	SPISVY	CASNE		Onvee	INTOR	NG	TTN		WN A
661			ERVVESOT	TETEL	DLUCLN	LDCIT.C	CILA TC	SPISVU	CICCO	NKAAH	P D V C C		NC	TATING		WNT.
657			TERESSO	REANT.	SLICED		DIASC	S DT S WW	GVCC	TRACC	A DIVCO		NG	TTM	7CCV DV	WGT.
056			DIVIDENCE - C	DUMPENC	VAUVAN		My Amc				CDMCC		C		COMPTR	UTOT
900	G - AP		DUALGH9	DIRKEVG	NAMAN	V D - FDT	HI GIG	GADANPV	a 1992	T T O O O	D F VIC CI				IGG TETP	

Supplementary Figure 6L. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 6** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30	40	50	60	70	80	90	100	110	120	130	140	1	50 157
Consensus	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	******	XXXXXXXXXXXX	XXXXXX SXXX	-XXX CCXX	*****************		xxxx	XXXXXXXX		-XX SXXJXXXX	-xxxxxxxx	XXXX – X 💽 XX	XXXXX OF XX -	-XX CXX	JXXXX
Identity														_	_	
1062	I PH P PK EYGDNI	WKPTGKKEWSD	KGHNGGDNGGN	GGGT SL SNG	-GGH	SVTSLADMAENTIC	-SLTSLEG-	IPLI	IDENLNVG		LG GV GG	GN		VHFD CHAN	GGSP	TNTL
1043		A PSGN		TM GNO	TPH CON	EVTEAONVGKHDVG	PL SGLEGS	SLGLVD TVL	GVTGNVG		TGOSVIG	NN		WTFK CIAVN	GOSP	IN IT.
966			A PO		-ST.V COD	ATSSGN-S.	TYMRAUGS	SALKMDI	VDDKMVA		TGGSSANP	-VNVGGGVS	SSTP-MORE	NSEG IC	TODAS	PVNV
959			0.55	SSN STO DDA	- PLT. CCT	YSGAGG-S.	DEMRANCS	SATHME	DPAKMYA		TGESEVNP	TNVGGGTE	TAAP-MOOE	NY FGT. WG	-JGCTT	PVGV
892			CLTHHT	TRGM SO GLG -	-TVO CE	GETESKDSEGA-Y-	-ALMOVIG	SVD	LSKKVLA		TGGSSSG	TGTLNGNT	KSSP-MOOD	NSEG CIAVE -	GOND	GSTS
908			TPAYER	DDG SN POTTT	-PLO CE	S-THTTD SALG-K.	-ALLOL SG	HED!	TGDGDTG		TGASPIG	GNT	TEOP-NON	-VTGTFIT-	GOVP	INT
996		SSLAPGL	VERTSHIKSAS	SGNVNSONVG	-FAO CO	TYOSHDORVO	SLGALLG	VVT-	P-MDGT.MAG		VO SPITN TL.	PILGGTST	RAOP-MOOT	NEVY	TOOSP	TST.
975		00000000	V PRTDR V PPVD	AAG	-GIY COT	SVVGGSDHAVR	-L TA ELLG	LAWG	PGTKGLVG		TT OSPITEAG	-PGGGNFC	SOOT - MCCT	DNFSCHUV-	- DOSP	NVNV
983	ANY ETNA SELARGI	PPTPTTRRTSS	WGAKRAOPSGV	PSN TSG NSG	-OVO COS	TVMN SAD SNAN	-TIMDIN.	MASP	-VTNL TG		I.N OSPI	-GGTGG -NR	TSOT -	IN SSD CIAVIS	TGOTP	SEROCIDES
742	ANY ETNA SRLARGI	PPLPPTRRTSS	WGAKRAOPSGV	PSN I SG N SG	-ovo	TVMN SAD SNAN	-TLMDLMN-	MASP			LNOSPI	-GGTGG -NR	TSOT - TOOLET	IN SSD CITYS	-TGOTP	SRROCHTS
709			LPLLA	LAATNO	-SAY CON	SVONATDAHA S	-NILEELG	TAVD	-VTGOVG		FG SPISV	VGVGAGSS	KOEP-MOOT	NSFN CATIN	GGSP	-LSL
675			NPAPT	EIEYEOREGG	-TVO	SYOKATDLNA EWT -	-KWLGFLN-	INAR	-VDANVG		FT CTGVKA	-GGIGGAAS	TOOK - THE	SN FN CVVA -	- IGOT P	L SAL
1004			TP-LVAR	-TEPTCS-ST	-EIK COE	NVGTAAVVSSVL	TPILVS-	TPLV	PLLSLLGLNA	VVALLEGT	DVVLGTCSG	IDVGG-SC	NAOT - MORE	VOFN CITIN -	-VGOVA	DIL
1003			TPAMLAR	GGGSSCAPST	-SLK	HVGTFSEVS	PY IN P-	LELV	V TAALLGL	-VGGLF	SVTLATCSG	ID IGG -SC	NSOT -	VVFN CHVN-	-7GOTA	ID IL
841			APSPGKG	GDTANV OGG	-TTY CONI	NFNO PVDESTG		VNVI	NLN			IFLGLGO	ASST-VEC	VGTN CLIN-	-VOOVG	AL
839			APAPGKG	GNTTNV COGG	-KTY CONI	NFNO PVDKASG	NLLSVP-	INVI	NLO		PISVLS	IVL PLGO	ASST-MOOD	VGTN CLVN-	-JOCVG	AL
840			APSPG-G	GGTANA OOGG	-ETW CONI	NYNO PVDN STG	NPLSVP-	EDVII	SAN			ILLGT-0	ASST-MOOD	MNT SCITVN -	-TOCIG	AL
837			AATDG	NT ONGG	-DTY CON	SSPLLVNPVTG	SFLAVP-	VSVL	TLV		CTPUTILG	IPIGG-T	SRTT-VCC	VDPT CUVN-	-70CIG	AL
423			APANHIL	SARDGCDANS	-OAH	TIANTTOT SAA SAI	-GKIGLSP-	EEAK	PVG		LT PTTLG	EVIDMGT-R	NSNA-700	NTYN CLIN -	-FGOVP	NINT
998	TP	VEVRDYOOFSR	VSLINFISFNK	ARGGDL ANG	KKTM CCG	ELKTFDKNILGLLI	GLLTTVVG-	LAGEI	KPTGLVG		LEVPITLLG	NGGL SDV	KANG - MOCKI	LTLRAA N-	-JGCVP	INTL
242			V PV EV ERKG P	SVDNKO SAG	T SVH CCD	TVD STDN SNVL	NOLNAAG-	IDH SDA	KOKGOVG		LTCTPITTSL	-IDA INGNV	OGAVTA CCE	T SOT CLVNL	NGCTI	TPVTL
927			S	PTRRDV GTR	-SLSCON	TIQ PAEQSSGA	FDG-	ELHVA	DLAIPVG		VT CT PV GALG	VATAGT	NTQS-ACCS	QT SG CLAN -	SP	INTGL

Supplementary Figure 6M. Sequence alignment and corresponding consensus sequence of proteins contained in Region 7 (full sequence)

	1 1	10	20	30	40	50	60	70	80	90	100	110 113
Consensus	C X – X X X X X C	C X X X X X •		X X X X X X X X X X	x x x x x x x	(X X X X X X X 🖸	***********	(X 🖸 X - X	x x x <mark>cc</mark>	GLXXXX	XXJXXXX <mark>LIS</mark>
Identity					_		_	-				
1062	S – NGGGH	CNSVTS		LADMAENTI	G S LT S L	G	-IPLNDLNI	LNVGLGCGVLG	3 N C O - O	O A VCC N N V H F D	GLINL GO	SPINIL
1043	G - NQTPHC	CNEVTE-		AQNVGKHDV	G P L S G 🖬	GS-LGL-V	DIVLNGVTO	GN VGLGCSVL G N	I N CQ - Q	O A VCC D N V T F K	GLVNL – – GO	SPINIL
1004	CS-STEIKC	CENVGT	AAVVSSVLTI	PILVSIPLVGPL	L S LL G L	AVVA-L	LFGTVDVVI	LGUTCSGIDVG		QTVCCENVQFN	GLINV GO	VAIDIL
1003	CAPSTSLKC	CDHVGTI	FSEVS	PYINPLELVGVI	A A LL G L	VG-G	LFTSVTI	LAUTCSGIDIG		QTVCCENVVFN	GLVNV – – GO	TAIDIL
841	CQ - GGTTY	CNNF		NQPVDEST	G – – – – D LL S VI	VN-V	IANLNCS	S P N V L A I F L G I	G – – Q 🖸 A – 🖥	STVCCQDVGTN	GLINVQC	VGLAL
839	CQ - GGKTYC	CNNF		NQPVDKAS	G N LL S VI	IN-V	IANLQC	PISVLSIVLPI	G Q 🖸 A - S	STVCCQDVGTN	GLVNVQC	VGLAL
840	CQ - GGETWC	C NNY		NQPVDNST	G N P L S VI	UD-V	ILSANCE	SPESVFAILLG1	Q 🖸 A - 5	STVCCQNMNTS	GLVNIQ	IGLAL
837	N - G G D T Y	CNSS		PLLVNPVT	G S FLAVI	? VS - V	LATLVC	PDT ILGIPI G (G T C S - R	T T V CC Q D V D P T	GLVNVQ	IGUAL
423	CDANSQAH	CETIAN			G K I G L	5 P E - E AK	- GPVGLT	PITT-LGEVII	MGTR CN - S	N A VCC S N N T Y N	GLINFG	VPINLNIL
242	CSAGTSVH	CDTVDS		TDNSNVLNQL	N AAG I	HS-DAK-Q	KGQVGLT	PITTSLIDAIN	I – G N V 🖸 Q G 🗛	V T A CCENTSQT	GLVNLNLG	TIIPVTL
998	ANGKKTM	CGELK			DGLLTTVVG	GE-ILK-P	TGLVGLEC	PITLLGNGGLS	5 – – DV 🕑 K – A	NGVCCKDLTLR	AAINVG	VPINIL
983	N - SGQVQ	CSTVMN-		SADSNAN-	T LM D 61	NM-ASP-V	TNLIGLN	5 P I G G T G G - 1	I R C T - S	Q T I CC E NN S S D	GLVSTG	TPSRRQCLIS
742	N - SGQVQC	CSTVMN-		SADSNAN-	T LM D 🗗	INM-ASP-V	TNLIGLN	5 P I G G T G G - 1	I R CT - S	Q T LCCENNSSD	GLVSTG	TPSRRQCLIS
709	CE-TGSAYC	CNSVQN-		ATDAHAS-	NILEEI	GI-AVDGV	TGOVGFG	PI SVVGVGA G S	3 S CK - Q	EPWCCTKNSFN	GAINEGO	SP-LSL
675	E - GGTVQC	CASYQK		ATDLNAEW	T K WL G F I	NI-NARQV	DANVGFT	GWKAGGIGGA	A S CT - Q	OK ICC TNSNFN	GVVAIGC	TPILSAL
996	CN-VGEAQC	COTIYQ.		SHDQRVQSL	GALLG	IPMDGL	MAG VQ	PT NILPIL G	TS-TCR-A	O P WCC T G N E Y Y	GLINIG	SPISL
975	CA-AGGIYC	CISVVG		GSDHAVRLI	A E LL G	WGPGIKGL	VGLT	SPIIEAGPGO	GN-FCS-Q	OTVCCTGDNFS	GLIVIDC	SPINVNV
966	CD-PDSLYC	CD		ATSSGNST	Y M RA V G	ALKMDLVD	DKMYATG	SANPVNVGGG	/ S S - S	IPWCCEGNSFG	- FT G TGC	ASTPVNV
959	D-DAPLL	CT		YSGAGGSD	E MRAVGS	ALHMEVDP	AKMYATG	EYNPINVGG G	E T-A	A P VCCE DN Y F G	LLVGVGC	TTTPVGV
908	CT-TTPLOC	CE-ST		HTTDSALG	KALLON	G-HEDYIG	DGDIGIG	PLGG1		OPNCC NVTG	TFTTTGC	VPINI
892	L-GT-VO	CEGET-		FSKDSEGA	Y A LMOV	GISVDALS	KKVLATG	SSG-IGILNG		S P MCC D DN S F G	GIIVGII GC	NDRGSTST
927	G - TRSLSC	CNTIQP		AEQSSGAF	D G LL H V	DL	AIPVGVT	PVGALGVATAC	G T C N - T	QSACCSQTS	GGLAN	SPINTGL

Supplementary Figure 6N. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 7** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30	40	50	60	70	80	90	100	110	120	130	140 144
Consensus	x 3xxx xxxxx xxx	*****	******	(X X X X 3 X X X X X X	(X – X 10 K X X	XXXXXXXX	*****	******	-xxcxxxxx	XXXXXXX	CXXXXX	exx	- x x x x x x x	********
Identity	1 .											-	I	L 🔳
320	A PUNI, KOTEVI, DECV	SDSVGSAGT	TEDVETLSA				WAGVESLPC	A APSTA PUT TA	FOA SCKUN	CDHRPI	NULD TL	ē	-SAUS	DEPVA
286	A PISY DRTS DDV	SKDVCEA	LEDVP TV ST	THELSE			ATEATSLEC	VAPSTAPET TAG	-FEASCLVT	GKKHPI	WILDLS		avvs	ln.
432	T PADS NL		VERLETLAP			T	TSOVESSNC	VAPUTPPITT	-FOAACLLD	CNRVAL	NOT IS AA		- No. WI	KAPGT
425		FPH	EKRETTEAO	WK PV	LAGEL	KTPLDD	T IDP TLETL	G TNAAS TYOS TO		SCKSEP(TEANLL-		RATIC	SPEKSKSS
421			A SVOTRGS	PN PL.	LAPET	KTPI.DP	TYDE TLOTL	GUNATA TIGS 70	LECEVYDD	SCSTAP	STEAFLL-		70 TIG	s
260		LPO	ESNLEMERSOF	DTSENGPL	LEGTL	TTP===LDP	TVDP BLLAL	G TDAAN TYOS	LLCHGYEE	TS CPTKPC	TEANLL-		RALC	ADER
24		LPO	ESSLSVKRS	DTSEWGPL		TTPLDP	EVDP ELLTL	G IDAAN IVGS	-LLCHGYEA	SS CDTKPO	FEANLL-	PST	ALG	ADEKE
384		FPO	EPNLSVKRS	DAST WGPL		TTPLDP	EVDP ELLAL	G IDAAN IVGS 76	-LLCHGYEE	SS CDTTPO	TEANLL-	96T	ALG	ADEN
178		LPS	AAN-SEKRO	SDRIGPL-		TTPLDP	LVDP ILELV	G IDAAA IVGS 70	-LLCHAYDD	T CTSEPO	TEANLL-	96T	GLG	GPLE HGH
380			APALQARQN	PANOR	A-KONAL	VP	TELGGI	IVS	-LGC IDL IA	GSCAAQTI	SCOQNNN-	<u>-</u> -v	VNV	DVL
193			FATVNAQT	BAAN-	T-KECOL	QDPDNLNA	DALNLI	N INPNT LTGA 76	-LTCTSLVS	GSCNANAJ	CSGN NYN-	<u>-</u> L	- IVIG	rozov
38		F	RDDSSATGADE	GG-TEAVGS	I-SETTN	SSGSDI	LGNVLGGSC	LLDNVSLISS N	-SNC	PAGNTH	PSNQD	5 -T	KSN-VS	IPVSA
31		F	RDDSSATGADE	GG-T AVGS	I-SECTIN	SSGSDI	LGNVLGGSC	LLDNVSLISS N-	-SNC	PAGNTH	CPSNQD	5	N IN-VS	IPVSA
37		1	PDDNSATGASE	GS-TEATGA	V-ARTTN	SSNSDL	LGNVVGGSC	LLDNLSLLS-	-SNC	PAGNTH	PSNSD	5 -T	N IN-AQ	IP IS A
36		1	PDDNSATGASE	GS-TEATGA	V-ABCTTN	SSNSDL	LGNVVGGSC	LLONLS LLSS	-SNC	PAGNTH	CPSNSD	5 -T	N IN-AQ	IP IS A
35	VTAYPGSSSAFG	VGQDEHKHHS	SDDHSATGASE	GA-T AVGS	V-SECTD	SSGSDV	LGNVLGGSC	LVDNLSLISI	-soc	PGANTH	CPSNQD	5-T	N TH-AA	IPVAL
88	L PLAD NAKANAV	T AR	TDGGSGSGSG	BDN HQQ	A-VECN	GI	LGC	LVSLGG	-SNC	AGS-TY	CSTNDAP-	5 FL	DINLLN	VFLR
40		л	PHGSSGGNNPV	B SA QNN	0-V	GL	LSC	A VQV 3G	-SNC	NGN-AN	CNTE APT-	6 TL	NVALLN	VELL
226	APTTG DGAGD GT	NGGGSSGGGS	scccsccss	GGGT NS NOV	A-VES	GL	IGGILC	N IGVIG	-TGC	SSG-S1	COSNAEQ-	SL	NINLLN	VAL
937		LPV	LGGVSRITSLA	ARDT DG GAI	MLLSFL	ASADSP	DVAP ILAEL	G IVSSSLDGL 76	VTCYPLDG	SNEGQ CAGQPI	CVNC SLD-	<u>-</u> -v	AFE	TEA
389	V MGD KVSR	QTGTSDLDVC	EDDSTTLPVT	ISTLGSI	L CDT T	ANGLT N	LDCAPPTKD	LTAVPS VSSLTE-	FYTACIIH	GRQRYPH	CILT VNG-	•DS	YLCEDPFA	TTVV
372			SPLSTRQAS-	ISN KLE	L CAG		VTETTYLTC	E PVFPF PQSVE E-	LEYRCED-	RGSNP(COFT	6LV	Y SQG	TKPGA
18			AP ANSDP SQ	TT AQJ	N-SCISI	TNG	IL	N VNV P-	-ALCVPLVG	NCNNQAJ	CETNGIP-	DTHG	CSQGLLN	LTVOV
429		SPM	ENGLWARDNG-	AQ AQ GS1	I-QECOSV	TSAGD GGL	LSNL LGLNC	AE IP IP ILN INT	TPRCNGQLA	CCSANTGTAPI	DDDDSNS-	GSQGI	ILSNLA I	PNIIL
803		TA	TQNLAPRQSGI	IS ALQ TLMLE	PVQCOEV	ELFSV VSG	LIGP ILSDL	G IPLPG LSTV 7A-	-FDC QPTLL	GLGCSEDTI	SLVVPN-	Q-D	- IG IN	DPVOPILR
219 (EAS)		A	PAEVVPRATTI	GPNTSIDDY	KPY COSM	SGPAG SPG	LINL PVDL	SASLGC VVGV	-sQC	GASVI	CKDDVTN-	TGNS	F 🛛 I INAA N	VA
321	V SKRGATAKRN SNVC	PGIEYPRCCE	ANDDGVIVSCI	YPGSVDSMEI	FRNACEYM	NKKDR EKHEE G	HKHEHGKEH	H KHGED KEHQE G	HEHEEKRH	GKKYRDGKQPN	CRLKQTKM	LPDVSQLFDR	KZAYND	T D L K D

Supplementary Figure 60. Sequence alignment and corresponding consensus sequence of proteins contained in Region 8 (full sequence)

Concensus		20 30	40	50 60	70	80 90	100	110 120	130 135
Identity									
425	VKPYLCCGEL	KTPL	DDTIDP IL ETLG I NAA	SIVGSIGLDCKAYDD	S	CKSEP-Q	C TEA	-NLLCGTTAL	SPLKSKSS
421	C PNPLL CC PEL	KTPL	DPTVDPILOTLGVNAT	AIIGSVGLECEVYDD	S	CSTAP-K	CTEA	-FLLCGTVALC	- S
260	VGPLLCCGTL	TTPL	DPLVDPLLLALGIDAA	NIVGSIGLCHGYEE	TS	CPTKP-Q	СТЕА	-NLLCGTTALC	BADLK
24	CVGPLLCCGTL	TTPL	DPLVDPLLLTLGIDAA	NIVGSIGLCHGYEA	\S S	QDTKP-Q	CTEA	-NLLCGTTALC	GADLKE
384	CVGPLLCCGTL	TTPL	DPLVDPLLALGIDAA	NIVGSVGLLCHGYEE	SS	QDTTP-Q	CTEA	-NLLCGTTALC	ADFN
178	CIGPLLCCGSL	TTPL	DPLVDPILELVGIDAA	AIVGSVGLLCHAYDD)TT	QTSEP-Q	CTEA	-NLLCGTCGLC	GPLEHGH
937	CDGGADMLLCCSFL	ASAD	SPDVAPILAELGIVSS	SLDGLVGVTCYPLDG	SNEG	QCAGQP-A	CVNC	-SLDGVVAFF	TEA
380	CPAN-QHAKCCNAL	VP	TLLGGIIVS	VGLGCIDLIA	\GS	C AAQT-A	CSQQ	-NNNGVVVNV	/BDVL
193	CSAANTKCCQQL	QDPDNL	NADALNLLRLLNINPN	TLTGAVGLTCTSLVS	GS	CNANA-A	CSGN	-NYNGLUVLC	S∎TQ I QV
88	CDNH-QQAVCCNGI	LG		CLVS L L	GSN	CAGST-Y	CSTN	-DAPGTLIDIN	ILLNGVFER
40	CSAQ-NNQVCCNGL	LS		CAVQ V L	GSN	CNGNA-Y	CNTE	-APTGTLINV/	LLNCVKLL
226	CNSN-QVAVCCSGL	IGG		I LCNIG V L	GTG	CSSGS-Y	CDSN	-AEQCGLINL	ILLNGVAL
38	CAVG-SQISCCTTN	SSG	SDILGN VL GGSC L LDN	VSLISS L N	ISN	CPAGNTF	CPS	-NQDCTIKSN	I-VS IPVSA
31	CAVG-SQISCCTTN	SSG	SDILGN VL GGSC L LDN	VSLISS L N	ISN	CPAGNTF	CPS	-NQDGTININ	I-VSCIPVSA
37	CATG-AQVACCTTN	SSN	SDLLGN VV GGSC L LDN	L SLLS- L N	ISN	CPAGNTF	CPS	-NSDGTININ	I–AQ G I P I SA
36	CATG-AQVACCTTN	SSN	SDLLGN VV GGSC L LDN	L SLLSS L N	ISN	CPAGNTF	CPS	-NSDGTININ	I–AQ G IP I SA
35	CAVG-SQVSCCTTD	SSG	SDVLGN VL GGSC L VDN	L SLISI L N	ISQ	CPGANTF	CPS	-NQDGTINIF	I-AA I PVAL
329	CKLGKPVCCPVD	VAGVES	LPCAAPSTAPVT I AG-	FQASCKVN	IG D	HRP-R	CVL	-DILCSAVS	SQEPVA
286	CELSTALCCDSK	ATEATS	LKCVAPSTAPFT I AG-	FEASCLVT	GK	KHP-R	CVL	-DLSCUVVS	5 0 D
432	QGTPVCCSSDL	ISGVLS	SNCVAPVTPPITTLG-	FQAACLLD	GN	RVA-R	CTI	-SAAGMAVI	KAPGT
219 (EAS)	CSIDDYKPY <mark>CC</mark> QSM	SGP	-AGSPG LL NLIP V DLS	ASLGCVVGVI	GS	QCGASV-K	CKDDV	-TNTGNSFIIII	IAANOVA
389	CTLGSPLCCDTT	ANGL TN	ILDCAPPTKDLTAVPSV	SSLTEFYTACIIH	IGR	QRYP-K	CTL	-TVNGGDSYI	EDPFATTVV
372	GNKLPLCCGAG	VTETTY	LTCEPVFPFPQSVEE-	LEYRCEDR	RGS	NP-Q	CDF	TCIVYS	-QG TKPGA
429	CAQG-STIQCCQSVT	SAGD	GGLL SN LL GLNCAEI P	' I PILNI I N	TTP	KCNGQL-A	CSANTGTAPPDCDD	DSNSCGSQGIN	ILAI PNIIL
803	CTLMLEPVQCCQEVE	LFS-VV	SGLIGP IL SDLG I PLP	GLSTVVAFDCQPTLL	GLG	CSEDT-L	CSLV	-VPNQD	IQPVQPILRS
18	CTTA-QANSCCTSL	TNG	ILNVNVLP	ALCVP-LV	/GN	CNNQA-A	CETNGI	PDTHGCSQC	LLN CLTVQV
321	CPG-IEYPRCCEANDD	GVIVSCTYPGSVDSMED	FRNACEYMNKKDREKH	EEGHKHEHGKEHHKH	IGEDKEHQEGHEHE	EKRHGKKYRDGKQPM	CRLKQTKM	LPDVGQLFDRKVAYN	IDETDEKD

Supplementary Figure 6P. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 8** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30 40	50	60	70	80	90	100	110	120	130	140	150 160	170	180 182
Consensus	XERXXXXXXXXXXXXXXX	*****		X X X X X X X X X X X	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	************	*****	*****	XXXXXXXXXXX	X X CC X X X X		-xxxxxxxxx	*****	*****	XXXXXXXXXIXXX	XXXXXXI
Identity			المسيحيا			and the second second	_		_							
973		SPTGDIDST	OVGT-FY	HGVODAHTS	-TVTKIALL HI	I-DEGNVTGPVG	LSCRD	DIDS	S-SCTDT-OF	LCCTGEN		FLEGIIVA	GCSPISL			
592		TADGHSC	TTGO-IV	NSLTSADAVP	-DLFSVTGL A	T-OAG-LIGNVG	VRCSTLT	VIGE	GKSCIASOOF	LeetDNN		Y-NGLFSI	GCSPINAMLV			
866	ARKLYAPTRAATAKF	QTTSATPSADT	TTGQ-LLCC	DELTAADSP	-DATNALAS	V-TLTSSE-DVG	LGCS	TLTT	DGTCDGTF	ACCGGT		INGAVSI	GCEAVTLAQ			
861	APRKLYAPTRVSAAKF	QSPSGVVTATADT	CTTGQ-LLCC	DELTPADST	DATNALAS	V-TLTSSQ-DVG	LGCT	ALTT	DGTCDNTF	ACCDGT		LNGVVSI	SCTAVTLGTG	SE		
854	APRRLYAPTRVSAAKF	QSPSGTPSADT	CTTGQ-LLCC	GELTPADSA	DAVDALAS	V-TLTSAE-DVG	LSCE	VLTT	DGTCGGTA	ACCGGI		LSGLVSI	GCITVAL			
853	PERRPYAPTATYGAKE	QEPSGSTTAS	PESGTN-LLCC	GSVTTPDNP	EASGILVD DD	I-MLGDVT-LVG	LD CT	VLDP	ADTCTTTF	TCCTET		LAGIISI	IGCVNVNV			
424		APSGLEERQO	CLNP LLCC	PDLKTPLDP	TLDPILLT G GI	I-NATALEGSICI	LNGQ	AYTT	TCATAF	PRCCTEVN		LLGG LV AL	GC			
377	IPA	PGEGPSVSMAQQ	GAEKVVSCC	NSKE	-LKNSKSGAE I PI	IDV L SGECKNIP:	INILTIN	QLI	PINNFCSDTV	7 S C C S G E Q		ILFSH I	RDACDHEGMMV	VEDVWSFCEGAG	VRRSALSESR S WL	LGALVSS
183 (MPG1)	IPA	PGEGPSVSMAQQI	GAEKV V SCC	NSKE	LKNSKSGAE	IDV L SGECKNIP:	INILTIN	QLI	PINNFCSDTV	/SCCSGEQ		IG LV NI	Q	CT	PILS	
358	GAVA	DSQAIKKQTEG-H	ODIGNVS	$\mathbf{N} \mathbf{P} \mathbf{T} \mathbf{N} \mathbf{V} D$	DKSDGFLTNL	WGVIGSLVNGQG	TACAPVSLIDE1	LGILALVKDT	PDGPVCENVI	ACCPGEG		AQCVAI	GDGSGSGSGYS	D		
229	GAVS	DSQAIKTQTEG-1	ODIGNVS	NPTNED	DKTDGFLNNL	WGVIGSLVNGQG	SACAPISLIDEI	LGILALVKDT	PDGPVCENVI	IACCEPGQG		AQCVAL	GDGSGSGSGYS	D		-
261	TAMS	LSESVALGGAKDÇ	OOSSGD ISOO	DSKENL	SADGILGNL	KGALNGLIGNDN	AACAKTSALDDI	LNVLSQTEDT	DNGPSCKNII	IACCPEGTGKL	TVQFLVHCYRQ	RINGSPLSLK	TKSGGHIVEQE	ENVTRSLLCMGV	KRSLVE L GV LN IW	G
21	MS	LSESVALGGAKD		DSKENL	SADGILGNL	KGALNGLLGNDN	AACAKTSALDDI	LNVLSQTEDT	DNGPSCKNII	IACCPEGTGKL	TVQFLVHCYRQ	RINASPLSLK	TKSGGHIVEQE	ENVTRSLLCMGV	KRSLVE L GV LN IW	G
262		LDHAKS		NSEEET	KADGILGNL	GGLLNGVLGNSN	SPCAKTSLIDEL	LNILASIKDS	NSGPVCKNII	ACCPEGT		TTCTAL	DNSGN	DD		
177 (DewA)	TATALPASAAKNAKLA	TSAAFAKQAEGTI		NSPAET	NNDSLLSGL G GA	AGLENGESGNTG	SACAKASLIDQI	LGLLALVDHT	EEGPVCKNIV	ACCPEGIT		TNC MAN	DNAGAGTE	AB		
105	UDDOVOWI	TSAAFAKQAEGTI		OCREPENCE CD	NNDSLLSGL <u>U</u> GA	AGLENGESGNTG	SACAKASLIDQI	LGLLALVDHT	EEGPVCKNIV	A GOPEGT		TOT DOM AN	DNAGAGTE	AB		
224	VPRGIQWIBHNDP	CECCLESSERVN1		QGFSPENGL-SK		EGVPGE DOGGNN		N	EEGARSV	ACCSEGS	N	NUTCIPUTAL	GRLM			
224	A LÖNIÖMT BHNDL	MDKCIWACCOCK		QGr SPEngL-SK		CCCCAVNUDALCI	INCET	N	DQG5K5V	VACESEGS	1	NETSEPCIAL	INACCURTING.			
380		APATOAPON				DTLECCTVSVC	COTD		ACSCARO7			NNCVVVN	IVCDVI.			
34	MPOLOAPOOAVE	HOGNAAVDAOONS	R POCYSDNTV PC		-VRPNCOTVSCI	VOVTSCSCCCAC	CVDHPD		FTCPGLOAVE			T.CVVS &	SCKNDTNTDH	KDSENEDCAODC	KTAOCCET.PET.AC	VSVACND
330	APA	ELVSRTTPVTPST	AGDYKPY	NSE	-PPTGLLPONT	AGLIG-WNVNAL	SCTVG	VT	GGOCSNNV	K CONGNAVN-		NGNSLIFT	GNVOCLL		and a second second	
294	TPI	AGGNDNAGDHSVO	GDOG-OTLY	NDISDSE	TKGEYYVGL	NVTVKCNDVTSN	VOSVDG	V	SROSOCDAKA	A GOSSSN		-TEONGETNT	VEGCOATAAN			
207		APGOWDAAAPI	OTDGGOLHO	OATETGAAAPVT	LAADLACYD	PAVENCWITDAP	IDPDFGCVGEYA	ACCOVNELNP	LLGLFCSKPF	GD RGREGG-	S	PAYCTDVIDG	REGNCTONDVE	NNOGLLGIGDGL	LGGGSK	
427	VPYAES	SGDGWGNGGDGGI	GGDGGNGGGGGG	GGGGGGGGGGGGGGG	GGGGGGGGSANYI	EACPSGLIGSYP	CCTVDILN-VI	LDLNCKPVPI	VPTSAAEFEA	AGASVGG	VP	ECCTVPLATI	GVLCEVPPGTA			
304	IPS	QISNDIEKRQSQE	CTTVGFTNPF CC	IP	-QLDVLNAVYV	ASECVQVPGGTS	ITSIQD	FD	SACLELNKIA	LCCNLNI		EG I ATC	SEAVQH			

Supplementary Figure 6Q. Sequence alignment and corresponding consensus sequence of proteins contained in Region 9 (full sequence)



Supplementary Figure 6R. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 9** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1	10	20	30	40	50	60	70	80	90	100	110	120	130 134
Consensus	VVAHPG.	******	****	XXXCCXXX	XBCCNKXXXX	x x x = x x x x x x x	X AX GX AX X	*******	XXXCSREXXX	x x x x x x - x x	XINXXCXXX	ACCOXXXXX	x ex – x vx x x	IX BCX X X X X III
Identity						a and	and a			and the second second	┕┛┖╴┛╸╸┙	المعرجة الألاء	معتور البا	all said
201		GRO TRA	PSTDOTTVACA	NN A CO NNM	VTCCN SUDN	PAGNAV	NGA GINGN-		EDOCSKIDVN	ZI. NTANG	LONG ROOM NO	ANCCONSERST	N Ro- চেল্ল সং	A REAL AND THE STATE
155		GPO TRA		NNA CONNMO	VTCCN KVTNT	P====AGNAV	NGAGUN-	N IS	P DO OSKUDVN	Zr. NTANG	T. ONE BOOM NO	A ACCONSERS	17 EG - 112N 72	
407		GPT TRA	PSTDOTTVACA	NNA CONNMO	VTCCNKVTNT	PAGNAV	NGAGINA	N IS	P DO CSKUD V N	Zr. NTANG	T. TANK R. CO.A. N.	ANCCONSERS	T BC - TRN 72	A REGIN LIS SING
383		LPSSSAA	TTPHENEVEDA	VNRCCDCAR	MSCONSINKN	EGT == SONNIT		-GSEG C	BOGCTOTOTE	ZEVICAAGLED	FARENCODS	ACCONENSO :	NDN-120	A 20VSFGDIAL
23		LPGSGHTA	HS-DNEVENA	VNORCNGAC	MSCCNTFHON	KNITOHNV	ZESNATEA		GOGCSOFNIP	ZSVAGAAGLE	FARENCODN	ACCONS SED	NNN-LING R	ALLOVSFGSION
255		APGGNY	GAGPREEGPSI	SROED-AC	TSCONKO IST	VNGGDKNS		TOXOGS 7G	E DO OS MIS	SMAIGVT	NAVEORED	ACCOG-DEK	No. 21	A REAL PLAN
159		APGGNY	GAGPREEGPS	SROED-AC	TSCCNEH ISP	VNGGDKNS		TICKCCS 7C	PDOCS KUS	SMIIGVT	NANSOCKOT	CCOG-DEK	NE 2 11700 72	FROM TAAM
160		APGGNY	GAGPREEGPSI	SROCO-AC	ISCONKO IST	VNGGDKNS		ICKCCS 7C	P DO OSKUPS	SMIIGVT	NUNSHORDT	CCOG-DEK	NE 0 1170. 71	TROPP TAAM
163		APGGNYI	GAGPREEGPSI	SROCO-AC	ISCONKO IST	VNGGDKNS		IGKGGS 7R	P DO OSKUPS	SALIGVE	NUNBOCKDE	ACCOG-DEK	AU	FROMP IAAD
164		APGGNYI	GAGPREEGPS	SROCO-AC	ISCONKO IST	VNGGDKNS		IGKGGSZG	P DO CISKUPS	SNLIGVT	NENSOCKOT	ACCOG-DEK		FPCTP IAAL
162		APGGNYI	GAGPREEGPS	SROCO-AC	ISCONKO IST	VNGGDKNS		ICKCCS 7C	P DO OSKUB	SNDIGVT	NUNBOCKET	ACCOG-DEK	e I IVA I:	FROM INVE
161		APGGNYI	GAGPREEGPS	SROCO-AC	ISCONKO IST	VNGGDKNS		VIITGGIVG	F DO SKIPS	SNIIGVT	NUSCENT	ACCOG-DEK		FPCTP TAAL
185		APSY	QGTPSVTKEQA	IARCON-SN	ISCCN NOV NK	ISAKGPDTDL	V NG MAKN	NLED IP	FDOCSKIN	AMEA	LINROCODT	AACCONTESN		VPCIPINIE
175		APSYI	QGTPSVTKEQA	IARCON-SN	ISCCN NO VNK	ISAKGPDTDI	V NG MAKN	NLED IP	F DQ CS KIN	INDEA	LINROCODT	ACCONTESN	N 2 V VA 3/	VPCIPINIE
209		APNY	AS-KPAVN-QKI	AAQCON-EN	LSCCN RO INK	VDATGADTDF	V FNG IVKN	N IQD IS	FDQCSRIN IP	CV CAG-LS	FINDRCKOT	A ACCONVNSE	N E A VA 77	VPCIPLNVE
272	VVAHPG	AGYVSTPEI	EANFQONFORF	VTACCION	VSCCNTETNK	GAPLTA	G FIP TONN	DLSD S	LKGCSKVD	VANVIGVQ	LONSNCRTO	SCCRTGDENO	VELINAN	KCAAHNIK
71	HPG	AGYVSTPEI	KEANFQONFORF	VTACCNGN	VSCCNTETKK	GAPLTA	G 🛛 I P 💷 D N	NLDDFS	LKGCSKVD	VANVIGVQ	LINSNCKTO	SCCKVGDENO	VOLVNAN	FREVAONIE
276	HPG	TGYSPE	ERNFRANFOKF	VTACGNNN	VSCCNED IKK	VGS-QSS	G 🖬 L 🖸 VIL D N	DLSDFS	LKG CSKL D	VAAVIGVQ	L L N S N CK T Q	SCCKSGDENO	a 🛛 Linan 🖉 🤇	ECAVONVE

Supplementary Figure 6S. Sequence alignment and corresponding consensus sequence of proteins contained in Region 10 (full sequence)

	1	10	20	30	40	50	60	70	80	90	100 105
Consensus	ССХХХ	O X SCCNKXX	X X X X X – X X X X X X	GXLXGXL	X X X X X X X X X X X X J	XXXCSKLXX	XXXXXXX-X	XXDXLNXXCXXX	XACCOXXXXX	XGX – XVXXX	XXPCXXXXX
Identity											
407	CCNNM	NTCCNKVT	NTP ACNAV	CNCACT	N N N S T	FDOCSKIDV	NULANC		ACCONSCCS		
155	CCNNM	VTCCNKVT	NTP AGNAV	CNGAGIL		FDOCSKIDV	NVLATANG -	L TNK BCOAN	AACCONSCCS		ALPCTALSST.T
201	CGNNM	VTCCNKVT	NTP AGNAV	GNGAGIL		FDOCSKLDV	NVLATANG -	I. T.N.K. ECOAN	AACCONSGGSA		ALPCTALSSLT
383	CCDCA	HMSCCNKTN	KNEGT - SONNT	GVIISNVI	GAD GSEGUG	GOGCTOIDT	PVSVIGAAGI	TTDETKKNCOOS	VACCONENSOA		ADCUSEGDIL
23	CCNGA	MSCCNTEH	ONKNT - TOHNY	GVISNVI	GAM = -GGEGHGT	GOGCSOFNT	PVSVIGAAG	I. T. D. F. I. K. K. N. C. O. O. N.	VACCONSSSOA		ADCUSEGSLL
255	CCO - AC	TSCCNKOT	STVN GGDKNS	GT. I. N G V L	GTWIGKGGSWG	FDOCSKILS-	-ISALIG	VTDNDNSOCKOT	VACCOG - DSKA		NIPCTPTAAL
159	CCO - A	TSCCNKHT	STVN GGDKNS	GLUNGVI	GTVIGKGGSVG	FDOCSKIS-	- ISATIG	VTDNINSOCKOT	VACCOG - DSKA	EG I.VA	NT. PCTPTAAT
160	CCO-AC	TSCCNKOT	STVN GGDKNS	GLUNGVL	GTVIGKGGSVG	FDOCSKIS-	- ISALIG	VTDNINSHCKOT	VACCOG - DSKA	EG LVA TI	NEPCTPTAAL
163	CCO-AC	TSCCNKOT	STVN GGDKNS	GI.I.N GVI	GTWIGKGGSWB	FDOCSKIS-	-ISALIG	VTDNTNSOCKOT	VACCOG - DSKA	EG T.VA	NT. PCTPTAAT
164	CCO-A	TSCCNKOT	STVN GGDKNS	GLUNGVI	GTITGKGGSVG	FDOCSKIS-	- ISATIG	VTDNINSOCKOT	VACCOG - DSKA	EG T.VAT	NT. PCTPTAAT
162	CCO-AC	TSCCNKOT	STVN GGDKNS	GLUNGVI	GTVIGKGGSVG	FDOCSKLS-	- ISALIG	VTDNINSOCKHT	VACCOG - DSKA	EGTVAT	NLPCTTTAVL
161	CCO - A	TSCCNKOT	STVN GGDKNS	GI.I.N GVI	CTWITTGGIWG	FDOCSKIS-	- ISATIG	VTDNTNSOCKOT	VACCOG - DSKA	EGTWAT	NT. PCTPTAAT
185	CGN - SI	TSCCNKOV	NKTSAKGPDTDI	GVINGLL	KNUNLED P	FDOCSKUN-	-TAATGA	TDLUNBOCOOT	AACCONTESNA		AVPCTPINIT
175	CGN - SI	TSCCNKOV	NKISAKGPDTDI	GVINGLI	KNVNLED IP	FDOCSKLN-	-TAALGA	- TDLLNROCOOT	AACCONTESNA		AVPCTPINI
209	CGN-EI	ILSCONKOT	NKVDATGADTDF	GVINGTV	KNINTOD IS	FDOCSKINT	PTGVTGAG-	LSDETNOKCKOT	AACCONVNSEA		AVPCTPLNVL
272	CGNGNO	VSCCNTET	NKVG APLTA	GGITTPTT	NNIDLSD IISI	LKGCSKVD-	-VAAVIG	VODLENSNCKTO	VSCCKTGDINO	VELINANVI	KCAAHNI
71	CGNGN	VSCCNTET	KKVGAPLTA	GGILIPIT	DNUNLDD FS	LKGCSKVD-	-VAAVIG	VODLENSNCKTO	VSCCKVGDANO	VELVNANVI	KCWAONI
276	CGNNN	VSCCNEDI	K K V G S - Q S S	GGLLGVL	DNIDLSDFS	LKG CSKLD -	-VAAVIG	VQDLLNSNCKTQ	V SCCK SGDTNQ.	AGLINANVO	GECAVQNVL

Supplementary Figure 6T. Sequence alignment and corresponding consensus sequence of proteins contained in Region 10 (without N terminus region preceding first cysteine of the eight-cysteine pattern)



Supplementary Figure 6U. Sequence alignment and corresponding consensus sequence of proteins contained in Region 11 (full sequence)

	1	10	20	30	40	50	60	70	80	90	100 106
Consensus	ССВХХОЦХ	CCNXXXX	XCDXIXXXXCX	XXCXLXBXJC	- X C X C X C J C	LXDXCXXL	* * * * * * * * * * * * * * * * * * * *	X X B J J 🛯 X X C X 🖸	XXACCOXXX	X X X X X X 🖬 X G X I	XLPCXALGSJJ
Identity			والمرجر الراقم								
241	CGNNNOVA	CCDDTTF	TEDOVEVESEP	LAGALKELIG	- G K N C A K G L G	LEDKCSNL		TDLISNHCKO	NIACCOGNI	ADSSSDUVGL	I P C V A L G S L V
240	CGTDNOIA	CCDDTTF	TODOVEVASOP	LAGALKDLLG	- GKNGAKGLG	LFDKCSKL	NII DVLIG	ISDLINSOCKO	NIACCOGNI	ADSSGDIGL	VPCVALGSLL
379	CGDQAQLI	CCNKTVK	TGDFTQVEEGL	LAGLLSNLLG	- AGQGSQGLG	LLDECTNI	PV-IPIIPIASI	PQ EQCKO	PISCCONTK	SADGDLIGV	ELPCIALGSLL
266	CGDQAQLI	CCNKTVK	TGDFTQVEEGL	LAGLLSNLLG	- A C Q C S Q G L C	LLDECTNI	PV-IPIISIASI	PQ EKCKO	PISCCONTK	SSADGDLVGI	GLPCIALGSLL
343	CGNDAVVS	CCNKQTV	SGSVINVNNGL	LAGVLSNAIA	- G C P C S D C L C	LFDGCSDL	SLGVPVLSLLS	AQNILNKKCQC	SIACCONSE	SEANGNLLGV/	LPCVALGSLI
361	CGSGNQLS	CCNKATY	AGDSTEIDSGL	LSGLLSNLIG	- SGSGSQGIG	LADDCSPLI	D 🖬 – – Q A L D L V G I	LQNLLNDQ C EQ	TAACCOGSO	SDTEGDLIGL(ELPCIALGSLL
230	CGDNAQLS	SCCNKATY	AGDTTDIDSGF	LSGVLSNLLG	– A G S G S Q G I G	LADECSPLI	D 🛛 – – Q A L D L V G I	LQNLLNDQ C EQ	TAACCOGSO	SDTEGNLIGV(ELPCLALGSLL
213	CGDNMQLS	CCNKETA	AGDTTDIPDGP	LADALKGVLG	GAGSGSEGLG	LLEQCGKLI	DII I P I I	ISDLINKHCDA	KAACCODSG	SSADNSLIGL	NLPCIALSLI
217	CGDQAQLS	CCNQASM	IGGDSINIDQGI	AAGLLKDVLG	– G G S V N Q G V N	LFHQCAKLI	DIQIP-IIGVP:	IQDALNQRCKÇ	N V A C C Q A P K	ADASNDGVAL	SLPCIALGSVI
191	CGDQAQLS	CCNKAIR	AGDKIDIG	GALKNIIG	- G G S G N Q G V F	IFDQCSHL	GVQIP-IIAIA	VQNILNEECHO	NVACCQTNF	GSASNDLGV	ELACIALGNLL
194	CGDEAQLS	SCCGRVVK	AGDRISVDEGI	GAGLLSNLAG	-G GSG ISS I L	AFDQCSRLI	DAQIA-VLIVP	VQNLLNQHCKQ	NVACCOKNE	GDSSSSGIAA	SLPCIALGSIA
180	CGDQAQLS	CCNDVKY	GGDTTVVQKGI	AAGLLSDLLG	- A C S A A E C L C	AFSGCSKLI	$\mathbf{D} \mathbf{I} \mathbf{S} \mathbf{L} \mathbf{L} \mathbf{I} \mathbf{G}$	IEDILNQK <mark>C</mark> KÇ	NIACCAKSG	GSADGDUVGA	LPCIALGSIL
422	CGDQATLS	CCNKAT Y	SGDTTNANSGI	LGGALSDLVG	- G G S G - E G L G	LFDQCSKL	PDIN-VIGVGI	LNDLVNQQCKÇ	NIACCQ-NS	S S E N T G U VAA (ELPCVALGSII

Supplementary Figure 6V. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 11** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	0 146
Consensus	LEPRO	x exxx x x x x x	ENGVGN KGN KBV	REXVEDBXTVK	AX X KCGD	QAQLS CCNKA	YAGD (TD IB)	COLLAX OF A	NL IGK GSGK	X GLGLF BOCSKILI	DLQ IP KJ IG IP	X ODDX NORC	KON IACCONSP	SK AKXX GI	KGLECK	ALGS IL
Identity		les en														
323	1	LOHDVNA	GNGVGN KGN NV	REPVEDD ITVK	AFEKCGD	QAQLS CCNKA	YAGDVTDID	SILAGEL	NL IGS CSG	BELGLFNQCSNVI	DLQ IP -V IG IP	IQALVNOKC	KON IACCONSP	SD AS BS L I	LGLPC IA	ALGS IL
189 (RodA)	1	L OHDVNA	GNGVGN KGN NV	RFEVEDDITVK	ABEKCGD	QAQLS CCNKA	YAGD/TD ID	6 116 A 6 8 16	NL IGS GSG	BGLGLFNQCSELI	DLQ IP-VIGIP	IQALVNOKC	KON IACCONSP	SD AS SS 6 1	LOLPCIN	ALGS IL
248	1	LEDHDVNA	GNGVGN KGN ANV	RFPVPDD ITVK	ABERCCD	QAQLS CCNKA	YAGD7TD ID	SILASE C	NLIGECSCH	BGLGLFNQCSELI	DLQ IP - IIG IP	IQDLVNQKC	KON IACCONSP	SDASSSIT	LGLPC IN	ALGS IL
195	1	LEDHDVNA	GNGVGN KGN NV	REPVEDD ITVK	AFERCOD	QAQLS CCNKA	YAGD/TD ID	3 3 11 A 6 E 1	NL IGE GSG 3	BIGLFNQCSEL	DEQ IP-IIA IA	VODLVNOKC	KON IACCONSP	SDASSSII	LGLPCIN	ALGS IL
216	1	L BO-DF NA	GNG F GN KGN ANV	RF IV PE BMT IKC	STERCOD	QAQLS CCNKA	YAGD TD ID		NL IGE GSGE	GLGLF DQCSKLI	DEQ IP-IIG IS	IQ DL INQQC	KON IACCOSS	AD ASCS 6 1	LGLPCVA	ALGS IL
176	L	PEAHDSOF	GNGVGN KGN 3NV	KF V PENVTVKO	ASDRCCD	QAQLS CCNKA	YAGDEREVD	3000 S 6A 0	GLIGAGSGA	BLGLFDQCSKLI	VAVL-IG	IODLANGKC	KON IACCONSP	SS ADGN 5 1	VGLPCVA	ALGS IL
385	1	AVGC VC	GNGVGNRGNDV	RESVEDNMTVKO	VD V KCCD	QAQLS CCNKA	YAGD TD IN	30111 C C F F	NLIGS GSGA	SGLGLFDQCSKLI	DEQ IF 7E IG IS	IQ DL INGKO	KON IACCONSP	SSASSD W	LGLPCVA	ALGS IL
19	LPP 5	A PSAGGAGS	GNGVGNKGNDV	RFSVPDNMTVK(ADAECCD	QAQLS CCNKA	YAGD TD IN	G III C G I I	NLIGS GSGA	SGLGLFDQCSKLI	DEQ IF VE IG IS	IQ DL INGKO	KON IACCONSP	SSANSDED	LGLPCVA	ALGS IL
26	DEB B(GASAGGAG	GNGVGNKGNBDV	REAVPONMTVRO	ADARCCD	QAQLS CCNKA	YAGD TD IN	GILCGL	NLIGS GSGA	SGLGLF DQCSKLI	DEQ IF 75 IGV	IQ DL INQKC	KON IACCONSP	SS AN SD DI	LGLPCVA	ALGS IL
265	LPP 6(S PNAGG AGN	ENGVGN KGN DV	REAVPONMTVRO	AD A KCGD	QAQLS CCNKA	YAGD TD IN	GML GGA F	NLIGS GSG N	GLGLFDQCSKLI	DEQ IP IIG IS	EQDEVNOKO	KON IACCONSP	SSANSDED	VGLPC IN	ALGS IL
20	DEB B(S PNAGG AGN	GNGVGN KGN HDV	REAVPONMTVRO	ADARCCD	QAQLS CCNKA	YAGD TD IN	GML G GA L	NLIGACSGA	GLGLF DQCSKLI	DEQ IP IIG I	FODFANGKO	KON IACCONSP	SS AN SD DI	SVGLPC IN	ALGS IL
346	L P P M	A PSAGG MG	GNGVGNKGNBOV	REPVPDNMTVKO	AD A SCCD	QAQLS CCNKA	YAGD TDVN		NLIGACSGA	GLGLF NQCSKL	DEQ IP-IIG IS	IQDLVNQKC	KON IACCONSP	SSANSDED	LGLPCVA	LGSII
25	1	AENMN	-NGLDHQNGGAP	REEVEDSLEVQO	NO SKCGD	QAQLS CCNKA	YAGD TD IN	36 III A 6E I	NLIGS GSGA	SGLGVFDQCSKLI	DEQ VP 7E IGV 9	VODLLNQQC	KON IACCANSP	SE ASND LV	AGLPCVA	ALGS IL

Supplementary Figure 6W. Sequence alignment and corresponding consensus sequence of proteins contained in Region 12 (full sequence)

	1 10	20	30	40	50	60	70	80	90	100 105
Consensus	CGDQAQLSCCNKA	YAGD XTDIB XGIL	X GTLX NLIG	GSGXXGLGLF	BQCSKLDLQII	XJIGIPXOD	LXNQKCKQNI	CCONSPSXAX	XXLIGXGI	PCXALGSIL
Identity										
26	CGDQAQLSCCNKAV	YAGDTTDINSGIL	GGTLSNLIGS	GSGASGLGLF	DQCSKLDLQIE	VLIGVPIQDI	LINQKCKQNI	CCONSPSSAN	SDLIGLGI	PCVALGSIL
19	CGDQAQLSCCNKAV	YAGD	GGTLSNLIGS	GSGASGLGLF	DQCSKLDLQII	VLIGIPIQDI	LINQKCKQNIA	CCONSPSSAN	SDLIGLGI	LPC VALGS IL
385	CGDQAQLSCCNKAV	YAGD	GGTLSNLIGS	GSGASGLGLF	DQCSKLDLQII	VLIGIPIQDI	LINQKCKQNI	CCONSPSSAS	SDLVGLGI	LPCVALGSIL
25	CGDQAQLSCCNKA	YAGD	AGTLSNLIGS	GSGASGLGVF	DQCSKLDLQVI	VLIGV G VQ DI	LINQQCKQNIA	CCANSPSTAS	NDLVGAGI	LPCVALGSIL
265	CGDQAQLSCCNKA	YAGD TTDINSGML	GGALSNLIGS	GSGANGLGLF	DQCSKLDLQIE	P-IIGIPLODI	LVNQKCKQNIA	CCQNSPSSAN	SDLIGVGI	LPCIALGSIL
20	CGDQAQLSCCNKA	YAGD TTDINSGML	GGALSNLIG	GSGANGLGLF	DQCSKLDLQIE	P-IIGIPLODI	LVNQKCKQNIA	CCQNSPSSAN	SDLIGVGI	LPCIALGSIL
346	CGDQAQLSCCNKA	YAGDTTDVNSGIL	GGTLSNLIG	GSGADGLGLF	NQCSKLDLQIE	P-IIGIPIQDI	LVNQKCKQNIA	CCQNSPSSAN	SDLIGLGI	LPCVALGS II
323	CGDQAQLSCCNKA	YAGD VTDIDEGIL	AGTLKNLIG	GSGTEGLGLF	NQCSNVDLQIE	P-VIGIPIQA	LVNQKCKQNIA	CCQNSPSDAS	GSLIGLGI	LPCIALGS IL
189 (RodA)	CGDQAQLSCCNKA	YAGD VTDIDEGIL	AGTLKNLIG	GSGTEGLGLF	NQCSKLDLQIE	P-VIGIPIQA	LVNQKCKQNIA	CCQNSPSDAS	GSLIGLGI	LPCIALGSIL
248	CGDQAQLSCCNKA	YAGD VTDIDEGIL	AGTLKNLIGO	GSGTEGLGLF	NQCSKLDLQIE	P-IIGIPIQDI	LVNQKCKQNI	CCQNSPSDAS	GSLIGLGI	IPCIALGSIL
195	CGDQAQLSCCNKA	YAGDVTDIDEGIL	AGTLKNLIGG	GSGSEGLGLF	NQCSKLDLQIE	P-IIAIAVQDI	LVNQKCKQNIA	CCQNSPSDAS	GSLIGLGI	PCIALGSIL
216	CGDQAQLSCCNKA	YAGDSTDIDSGIL	AGTLKNLIG	GSGHQGLGLF	DQCSKLDLQIE	P-IIGIPIQDI	LINQQCKQNI	CCOKSTADAS	GSLIGLGI	IPCVALGSIL
176	CGDQAQLSCCNKA	YAGD TTVDEGLL	SGALSGLIG	GSGAEGLGLF	DQCSKLDVAVI	,IGIQD	LVNQKCKQNIA	CCQNSPSSA	GNLIGVGI	LPC VALGS IL

Supplementary Figure X. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 12** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	1 10	20	30 40	50	60	70	80	90	100	0	110	120	130	140	150	160	170 177
Consensus	XXXXXXXXXXXXX		-XXXX EXXX	XXOCCATDVLGXA)	DCXXPXXX	EXXXXXEXXXCXX	XGOXAXCCVLP	K>	K X 🛛 X X 🗖 C X	XPXCXXXC	EXXXXXX	XXXXXXEEX	XEXXXXXXXXX	Xexxxxxxxxx x x x x x x x x x x x x x x	XXXXX	XXXXXXX	XXXXXXXXXXX
Identity	-		and the second			and the second						_	_	a seller ser			
433	PTC-NTCCCSS		- WVACSCEVIS	ST.OCCA TDVTCVA			A COPAR COVT D	I									
281		C SCCCC SCCCC SC		STOC CATDVICTAL	TROC DETEXA	PUCAN DIVATION	CODAD COVID			CDTANTY							
401	LDTFFKP()		VTPCSCIVC	TOCCA TOVICUA	C NPD C	DA NATING AVCON	TCOPAPCCVTP		DOCTON								
360	I DEFERROR		VTRCCCTVC	TSOCCATOVIGVA		TANAT DESAVOS	TCOPARCEVER.			TPICKOD							
303	I DA NEKD OA		VUDCmc VC	COCCATEVICUA		ANA TOBOAVOOR	TOODAD COULD		DAGTION								
254	I DA NEKROA		VTUCCCUVC	I SOCCATOVICY	NOC NED CO		TCOBAB COVIDE										
30 (EcHydSo)	I DA NEKROA		VIDCSCIVC	TSOCCATOVIGVA		ETDADNESAVCSE	TCOBARCOVER.			TPICKOD							
300	LDAGEKDOT			TSOCCATDVLGV			TCOPAPCCVTP	ACTT	DOCITICN								
402	A DI FEROA			TA OCCA TOTICUAL	TRACINERE	VTSADNIJSEV CSA	TCOPAPCCUTP	A 5 I D.		TDTAVSD							
303	A PLDER OA			TA OC CA TOVICYAL	TRUCK NEDET	VTSADNESEVCSA	TCOPARCCUTP		DOC TTCN	TDTEVSD							
431	A DI FEROAV		- WUDCSCI VC	USOC CA TOWL GUAL	TROC NEDEA	PTCADEDCCVCCA	TCODADCCVTD		DOCTON	TOTANTD							
390	VPLEEROTP			TSOCCA TDVIC TAL	DECONDRAT	VTSADNIJSEVCSA	TGORARCOVIP	i	DOC THON	TPSEVOD							
412	APLEEROT		YTECSGINA	SA OCCA TOVICIA	DEGDETAV	PTDAA NIJSA TCAA	TGORARCCVLP	i	DOGTOOD	TPACVDD							
328	APVIEEROVP		YICSGINA	SSOCCA TOVIC TAL	DOGTEPSA	PASADERSAVOS	TCORAR COVIDE	i	DOGTOCON	TPSEVOD							
258	APVVEDROV		YIRCSGING	SPOCCATDVLGVA	TDCGEPPAV	PTNASEBOATCAT	ICORAR COVID		DOGVICN	NPACVDE							
269	PAKEVESROVP		YIPCSGIVG	SAOCCATDVLGLVI	DCGOPPET	PTDADTESAICS	ICORAR CCALPY	7	DOGVICN	TPACVOP							
186	PA-EVOEROVP		YTECSGUYG	SAOCCATDINCIA	DOGODSDA	PV DADNISE ICAA	IGORAR CCV LP	j	DOGINCN	TPACVTP							
285	APVVEEROTP		YVPCSGIVG	SAOCCATDILG TAL	DCANEYEV	PTTADNESSICS	IGORARCCILP	j	DOG TICN	TPACVOD							
220 (NC2)	APAAMERQVP		YTPCSGIYG	TAOCCATOVICVAL	DCANEPAT	LANATHESTCA	IGQRAR CCV LP]	GODILCO	TPACL							
417	PTAEVHERGD		YTACTGING	TSOCCATOVICVAL	DCANEPEV	PISADNESEVCSA	IGORAR CCV LP]	GODVICO	TPVEVDS							
166	APSADANGVYRRA	N	AFCPDGLLYT	NPOCCDA DVLGLAI	LDCVTPPSA	PTSCKSEDGICAG	IGREPKCCAVP)	IGVALLCV	DPVPQ							
341	APLTGYSSGAS-		-SGVCPPG-LYS	NA QC CA T DV LG AA	LNCKAPSTT	PTSTDQHISG CAA	TGQQAKCCVLP	2	AGOADECQ	DVTPSANC	NNGGANGG	NNGGANGGN	NGDNGDNDD	DNN GGA YGAGA	ASQTPTS-1	PASOA TPC	PSSSPAAVPY
319	APLTDYP		-TALCPAG-LYS	NA QCCATDILGVA	LNCONPTTT	PTSTGDEISGCAA	VGQQAQCCVIP	72	AGOALLCQ	DVSPSGNC	ING DANG	NTGSANGGA	NGSADD	GANGATGGGA	AVQTPES-S	PASOA TPC	PSDAPN
338	APLASYS		-SGVCPIG-IVS	NPQCCATDVLGVA	INCONPATE	PTSADQEISG CAG	GQQAKCCVIP	71	ANOAVICO	DVSPNAKS	NGSANG	NGDNNDDG	NG-DDADGDD	DedDNHGAGD	SAKTPTSS	PASOE TPO	PSDVPY
332	APFVTTG		-TSLCPAG-INS	NPQCCATDVLGAI(LDCAVPATT	PADVDAFISGCA	VGQQAKCCVIP	2	AGODILCQ	DVKESGGE	GGSGGAP	I TA TGG CCG	GETPTSVPGG	GGGGATSTSC	SSAVVEPPI	POPTKPAPS	YPTTTSDCGC GK
233	APFAMTE		-TSLCPAG-LYS	NPQCCATDVLGAI(LDCAV PATT	PANVDAFISGCAA	GQQAKCCVIP	77	AGODILCQ	DVKCSGGC	GGAPG	TTATGG GGG	APTSVPGG	GGG-ATSTSC	SSAVVTPAI	PQPTKPAPS	YPTTTSDCGCGK
133	GPVEVRTGT		-APLCENG-INS	NPQCCDILVLGIV	GCEVETQT	PRDGVDEKNICA	TGDQAACCVLP	71	GOD HEC Q	TPVCIQG							
130	GPVEVRTGT		-APLCENG-EHS	NPQCCDILVLGIV(GCEVPTQT	PRDGVDEKNICA	TGDQAACCVLP	7	/GOD LLC Q	TPVEIQG							
132	GPVEVRTGT		-APLCLNG-INS	NPQCCDILVIGIV(GCEVPTQT	PR DGV DEKNVCA	(TGDQAACCVLP)	7	/GOD LLC Q	TPVCIQG							
145	GPVEVRTGG		-GSICEDG-LES	NPOCCDIQLLCII	GCEVESQT	PR DGA DEKNICA	TGDQALCCVLP	2	AG OD ILC Q	AAVCAA							
259	APATLHAR		-ATVCPTGLLYG	VAQCCATSVLGVA	DCSVPSST	PSDGADLKRICA	SCAAMCCSIP	2	AGOGVICT	PVIC							

Supplementary Figure 6Y Sequence alignment and corresponding consensus sequence of proteins contained in Region 13 (full sequence)

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	141
Consensus	CXX G	FYXXX DCC M	DVLCKAX DDC	XX EXXX EXX XXX	SXXX SXX	X OQX XX COVIDEX X	X X J CCX	X P X CX X X CCX X	XXXX	XX EGXX EXX XX	XXXX	* * * * * * * * * * * * * * * * * * * *	******	* *****	XXX
Identity				ليرون والمراجع								L	_		
20 (51) 15-2		_													
39 (FCHydSp)	CSG	TYST SOCCA	D VILC VAD DDC	GN PSSETDADN	IF SAV CAE	IGO RAR COVIDED	D OB THE N	TPT VQD							
354	cs -	TYST SOCON	D VIEC VAD DO	GN OPSSICTDADS		IGO RAR COV LP I	in die mie n	PPT WOD							
401	cso-	THET STOCK	DVLGLAD DDC	GN SPSTEAN ATT		CORARCOVIET	h da me s	PPT WOD							
80	an a	AVESSOCOM	D VIC VAN TOC	GT JPSV JAN ATT	BEAV NE	IGO BAR COVIDE TH	n de me s	PPT MOD							
303	CS	AVET A DOC M	D VICE AD THE	ON PETVTS ADS	EREV SA	TCORAR COV TR TH	b is men	TPT WSD							
402	CSG	TYST ACCA	DVEGVADEDO	GN PETVTS ADS	E SEV SA	IGO RAR COV LP II		TPTOVSD							
390	CS G	LYST SOCCAT	D VIC IAD DO	GN PATVTS ADS	F SEV CSA	IGQ RAR CCV LP I		TP S V QD							
431	CS G	LYSTSOCCAT	D VEGV AD EDC	GN SPEARTS ADE	SSV SA	IGO RAR COV LP I	D D B TILC N	TP T SV TD							
328	CS G	IYASSOCCM	D VIC IAN DO	GT PSAPASADE	F SAV SS A	IGQ RAR CCV LPL	D B TLC N	TP S V QD							
186	CS G	LYBSAQCCA	D HIGL AN DO	GQ SDA EVD ADS	IF SE I <mark>CA</mark> A	IGQ RAR CCV LP I	D D B IIIC N	TP AGV TP							
269	CS G	IY SS AQCCAT	DVLGL VNLDC	GORPETPIDADI	FSAICSA	IGO RAR CCALPV L	D B VLC N	PPASVOP							
285	CS G	LYBSAQCCA	D ILG IAN DO	AN SYEVETTADS	IF SSIC <mark>S</mark> E	IGQ RAR CC ILP I	D D B III C N	TPA VQD							
412	CSG	INAS A OCCAN	D AFCT VD TDC	GD STAVETD AAS	FSAICAA	IGO RAR COV LPIL	b ge me b	TPA VDD							
258	CSG	TYSSP DCC M	D VLGV AN LDC	GE SPAVETNASE	E DAT CAT	IGQ RAR CCV LP II	D D B VIC N	NPAOVDE							
220 (NC2)	CSG	TYSTAQCCA	D VILCY AD DOC	AN PATLANATH	EST CAN	IGO RAR COVIDID	e do me o	TPACL							
41/	OT C	THET SOCON	D VIEC VAD DO	AN SPEVILISADS	IS SEV ON	ICORAR COVIDER	ie go Ane Di	PPV VDS							
422	cso-	THUSL COM	D VLGLAN DO	SP STRASTS AND		O PAR COV LP 1	is ga ang pa	SPL OVIII							
130	Gen G.	THESE CON	r. and the res	EVETOTERDOVE	ISKNT ON K	TO DAA COULEV	is in Arc h	TRVETOG							
133	GPN G-	WENP DCCD		EVETOTERDOVE	EKN I WK	T CD DAN COVILEVY		TRVEIOG							
132	CLN G	WSNP DCC D		EVSTOTERDOVE		TODOAL COVLEWY		PVOIDG							
145	CPDG	FENPOCCO		EVESQTERDGAD	FRNICAK	TODOAL COVIDIA		AAV JAA							
233	CPA C	LYSNP DCCM	DVLGAIG DDC	AVEATTRANVDA	ISG CAA	OQ QAR COV IPVA	G D MCC	DVK SSGSGSG-		GGEGGAPTS	VPGCCCC	ISTSC SSAV	V PAPOP KR	ASYPTTISD	CGCCK
332	CPA G	LYSNPOCCM	DVLGAIG DO	AVEATTRADVDA	F ISG CAN	VEQ QAK COV IP I	GOD MCCO	DV K BS GG BG SG S	GGAP	GGGGGGGGTPTS	VPGG COC CC	ISTSC SSAV	VRPPPQPREP.	ASYPTTISD	CGCGK
259	CPT GE	LYBVAQCCAT	S VILOV AD DOO	SVESSTESDGAD	L KR I <mark>CA</mark> E	SGAAMCCS IPLA	GOS 71C II	PV IG							
319	CPA G	LYSNACCAT	D ILCV AA ENC	ONSTTTPISTGE	F ISC CAN	A 20 OVO COA IBAV	.G 01. 01.C 01	DVSPSGN SG I NG	DANG NTGS.	ANGGANGEN NG	SADDGGANGA	I GCGA ONVOI	PRS-SPA30A	T SCPSDAPN	
341	CPP G	LYBNAQCCM	DVLGAAA DNC	KASSTTETSTDO	H ISG CAA	DO OVE CCATE I	COV UTCO	DV TPS AN SGNNO	GANGENNEG	ANGGNNGCDNC	DNDDDNN GG	Y GAGA OA SOT	PRS-YPASOA	T PCPSS SPA A	VPY
338	CP I G	LYSNPOCCAT	D VIC V A DINO	ON PATTETS ADO	F ISG CAG	EQQARCCV IPVA	N OV AFC D	DVSPNAKSNC	SANGONNOD	NND DGN 🖸 – D DA	DGDDDDD	HGAGDOSART	PESSYPASOE	T CPSD VPY	
166	CPD GL	LYENPOCCDI	D VEGE AD EDG	VTPPSAPTSCKS	E DG ICAG	IGREPKCCAVPII	GAY PROA	DPVPQ							

Supplementary Figure 6Z. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 14** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

Conservus Sizta + 3 internet and internet a		1 1	10	20	30	40	50	60	70)	80	90	100		110	120	130	139
Identity Ave at LUMANES 0 Ave at	Consensus	X P TAGK Z PK	******	xxxxxx	*****	XXX XX XX X 3	х ссххв и х	хах в жхз	*****	F XX X CA XX	OKXXX CC	XX XX XC		XXIJ	******	XXXXX	SALQFLWHSC	GLAIAVH
30 No I I LANARD 1	Identity				— — — —													
30 AP = 11.48.07 = 0.73 AP = 1.48.07 = 10.21.07 AP = 11.48.17 = 0.74.07 AP = 11.47.17 = 0.74.07 AP = 11.47.		_																
1/1 UPA IT ALL COMMOND - UP	30			APSI	KLRAP SD-V	PALDT	LSCOVDAT	J. D. T. EAS	s-ddt sv sn	F BAACA PT	GL TAR CC	TLPLLC		EVPTCE	TP			
1/4 VPAIRADAUVARUS. MAGD. VB	317			IPAMI(DERAPGS-V	PALDT	L COVDADO	VDVTERS	S-DTS SV NA	EKSS CAS S	GT TAK CC	TL PL AG		DALLCS	SP			
32 32 Min Dis You Handle Act of a Act of a Yee - Alt Y a Hold B UP Hold B Y B I KN BY CY VD 2 11*0 To Act Of Wild Act CARDEY DL 387 Inck B UD - BOY	174		VPAIKRDA	EVPAMI	K-RAAGD-V	PALDT	MOCSTRUE	JAD SST3	S-DTS SVAA	NSAUNT	GK TAE CC	TVEVAC		OVINCE	SP			
315 TEXTS T	32	NPIAGNOPO	SYGISGACO	SGAGOS	SEGDAAG- I	TALYT	D COAS III O	ADDA TP3	SL SVN SKQT	LVDDCANT	GATAQ	AT 5EVC		OVPECA	DL			
38/ IPPER AND 0-0007	315	TPIAGEEP-	GISGAGT	YGPGQV	SDGDAAG- I	TALYT	D CETS 70	ADJACNTS	SVSVDSEQS	LINDAKT	GA TAE CO	AT 5AVC		0 V VICY	DL			
New protein seq IFRE Number Roy Control and Control and State a	387		IF	GKSALVI	D-RQVTL	SGADSSA	VCATOVE	ZADIDICAP :	PS TPT SI DD	STS ICASI	OD QAR CO	LL PILE		QG <u>FIC</u> F	SPGK			
72 AP bD = 80A b =	New protein seq		IF	GKSALVI	D-RQVTL	SGADSSA	VENTOVER	ADDICAP:	PSTPTSIDD	TSICASI	DOAR I	LL CILE-		OC UTCL	SPGK			
333 AP LD = REV - LD =	12			AP LD I	E-RQAGL	SSGN 5	VEENTDWEN	AN PLAP :	AT TPT SINA	IDG MAA	ODDAK IC	LINLG		QA PICE	DVNPTADI	TARRAL		
234 AF DO F. BOY CO. 1.2 Der. 1.00 (2.1.2.2.0.1.2.0.1.2.2.0.1.2.0.0.1.2.0.1.2.0.0.0.1.2.0.0.0.1.2.0.0.0.1.2.0.0.0.0	333			AP LD I	E-RQAGL	SSGN 5	V COAT DVIES	N N PDIEA P 5	AT TPT SINA	IDG MAG	OD DAK HE	LL DILG		QALLES	DVNPTAA			
350 VAX W = REQUEST = Star = Star = M = 24 A 2010 S = M = 24 A 2010 S = M = 24 A 2010 S = 000 H = 24 A 2010 S = 000 H = 24 A 2010 S = 24 A	234			AP LD I	E-RQVGL	SSGN 5	VEENTOWEN	AN PLAP :	AT TPT SINA	IDG <u>AA</u> G	ODOAK CO	LL HLG		QA PICE	DVNPTAA	a The states		
166 AF BEDARDOUT DEAL-TEDE AN CARTING AND AND AND AT TANK AND AT LEVEN AND AND AND AT LEVEN AND	350			VPAVVI	E-RQAGE	SSGT	V CONTINUE	TAN PROAPS	AITPTAINE	E ID T CA EG	ODOAK CO	LIG		QA HICE	DVNPTAE	PARAA		
310 AP UP = ROY G = UB =	188			APLET	DKRQVGL	ASGS	VERATOWES	AN PICAPS	TTTPADVNT	S ID V MAFG	ODOAK CO	LIG		QA nnes	DVNPTAA	PATPSA		
337 AF USE - REVICE 10 Sectors SIG 30 (2010) For an Sector SIG 30 (2	310			APLES	E-ROVGE	SSGN3	VEENTOVIC	AD PROAPS	RVTPAGVNE	SIDG MAD	ODOAK CO	LIGILG		QA HICE	DVN P	IP MA		
342 AF UR TRANSFORM AF UR TRANSF	337			APLES	E-ROVGL	SSGN3			TV TPAGVNE	J ID T SSV	SDOAK HE	LIGHLG-		ON HIGS	DVNPTAP	P DAAA		
320 AF UD A = UD A	342			AP LEI	E-ROVGL	SSGS3		A D D D D D D D D	SVIPAGUNE	TD T C T SQ	DOAX 10	T TOTAL		O D D TOC	DVNPTAP	D D A A		
315 ar the construction of the set of the	212			AP DO I				The Device A	STIPISIDE		DOAX 20	T T - Bot O		ONCO	DVAPEAR	a Smell		
90 AF DE RECONNECTOR SECONDATES DU DITES DUDITES DU DITES DU DITES DU DITES DU DITES DU DITES DUDITES DU DITES	515			APTER	-BOCDAVE	TOTNONA	O CONTRACTOR		RTTPRAVDE.		DOAM 20	TT DOLO		O N AVAR	DVSFanut	1.91.9		
AP TIPSROVC - LOAD - CONTROL - CONTR	352			MPTT	TSRODA T	99CN	b Conversion	TAD BUILDE AS	PAAVTDIKS	SEDV DUDV	ON THM CO.	DL DULC			SPRNC			
361 3	281			APTIL.	PORCHART	38GR			PRUPTOTOS		OD TNO CO.	TT DOLO			CDDCC			
362 PFAULT BLOCK - LOST - G - T 20 GOV 2020 TANDED IT PEGTPANTED DATE TO DEAL TO THE ADDRAL OF TANDED DATE 306 SP LOP - ROLG LOST - G T 20 GOV 2020 TANDED IT PEGTPANTED DATE IN TO DATE TO DEAL TO THE ADDRAL OF TANDED DATE 306 SP LOP - ROLG LOST - A LOTAL GOVS DA GENS TO TANDED Y SHELS RE LONAL DATE 98 APS SD UN GUIRE ANA F PEGLUYN 3L GOL 2020 TANDED Y DEAL STUPP TILT DA LOT GOAS GO VUIDA AND TANDED Y DEAL STUPP TILT DA LOT TA TANDED Y DEAL STUPP TINT DA LOT TA TANDED Y DEAL STUPP TILT DA L	373		S F	AN GA DU	NROLGL	SSAATN	D CONVERSION	A D D O V	PT TPT SUPD		OR THM CO.	TT. DOLG		O A GROT	SPDSET			
302 SPEC	362			PDAUL.	NDROVA T	00AAT 83	N CONVERSION	No. D. COLOR TOP 5	POTRANUTO	SNDT. SARV	TOP OF	TT DUCT	PPPCAPUP	VEUPTA	OPPNELOI	AVMENT	CALOPIN HOC	OT. A TAVH
900 SP DO F ARDON AL DO VALUE DO VALUE DO VALUE DO VALUE DA LANCE	306			SPLOT	P-ROLGL	SSA TOTA	O COSVS THE		HUPPT SUES	SRE TAAR	Shoas 20	A17 -37 A 6		O & VICOL	AASA	CAL WEED ND	DADGI DADDO	dire in vii
200 APS \$0 UNG ILERANA P PEGLETN & GOL AUTO CAN DUDO V PAXES \$CKS SOV SET SUPER CONTENDED AND TO PERAINANDED AND THE SET SUPERAINANDED AND THE SET SUPERAINDED AND T	428			SPLOT	P-RRSGL	SS==ALDTA	T. CODVS 7A C	TANK AND A S	STTPTTAT	DA TANG	BOAS 20	VT. DRAG		EARIER	AP			
APAG APAG SEG PGGLLENVIL GATINTO DATING TO STITUPVLTPAA DESENT KORV GG-STAAG LGATGENPIGTO 85 APQG SEG PGGVLENVIL GATINTO DATING TO STITUPVS TAAG DAE ASK KOPV GG-STAAG LGATGENPIGTO 139 APSG SEG PGGVLENVIL GATINTO TATING TO STITUPVS TAAG DAE ASK KOPV GG-TA JAG LGATGENPIGAO 140 APSG SEG PGGVLENVIL GATINTO TATING TA	98		7	PSSDUN	STIRRANAE	PEGLLYTN	T. CODI. DVIC	A DVDCVV S	PAKPSSCKS	SSV ST	REPR	AVENAG		VALLER	DPIPAI			
85 AP QG NEG S PGGVLENV SU GATING O AN TOG DE TEVP VS TA AQ DA E TAK KOPV GG - SLOVAG	84		-	AP	AGSEG	PGGLLHNV	L COATNEED	ATTOST	TV PVL TPAA	OSHAGK	KOPV C	SIAAG		LGARCE	KPIGTO			
139 APSG	85			AP	DGNEG	PGGVLHNV	LEGATIVE	AT FIGT 3	TV PVS TA AO	DAH	KOPV C	SL WAG		LGALCE	KPVGTO			
140 APSGKEG PGCVTNTVJ G GATNVIG NT JEG V TK PVP DVTA DA F NSK SKOPV GC-TA JAGLGJGEKPTGAQ 94 AV PTGLFSN - J GATNVIG JIG JEKT TI AVD TGA I DA HASK SKOPV GC-TA JAGOAJGO KA IGTF 97 AV PTGLFSN - J GATNVIG JIG JEKT TI AVD TGA I DA HASK SKOPV GC-TA JAGOAJGO KA IGTF 104 AV PTGLFSN - J GATNVIG JIG JEKT TI IAVD TGA I DA HASK SKOPV GC-TA JAGOAJGO KA IGTF 112 AV PTGLFSN - J GATNVIG JIG JEKT TI IAVD TGA I DA HASK SKOPV GC-TA JAGOT JGO KA IGTF 112 AV PTGLFSN - J GATNVIG JIG JEKT TI IAVD TGA I DA HASK SKOPV GC-TA JADOT JGO KA IGTF 112 AV PTGLFSN - J GATNVIG JIG VEKT TI IAVD TGA I DA HASK SKOPV GC-TA JADOT JGO KA IGTF 113 AP AT SPNGTD A PDGGL IGT D GOSL JEG VEKT TI SKA SKOPV GC LAG. SKAT DE JGA SA VEKT SKAT STOLKAR SOT SKAR SOT	139			AP	SGKEG	PGGVTNTV	LEGATIVE	ALC ODD OV S	TKPVPDVTA	DAH	KOPV C	TA PLAG		LGLCE	KPTGAO			
94 A VE PTGLES N. J. COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OA JECO KALGEF 97 A VE PTGLES N. JI COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OA JECO KALGEF 104 A VE PTGLES N. JI COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OA JECO KALGEF 112 A VE PTGLES N. JI COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OA JECO KALGEF 113 A VE PTGLES N. JI COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OF JECO KALGEF 114 A VE PTGLES N. JI COATING DIG JECKT JI LAVD TO ALE DASK OS KPLICE VALVAD. OF JECO KALGEF 1153 AP AT SPNG VE A PPOGELICE DISCUE ALE SE SEXTENSA KEE DE LICHT OF SE VERTE VALVAT. OF VERTE VALVAT. 298 LP AT SANG VE A PPOGELEGN DISCUE JESTICA JIK TEN SA KEE DE LICHT OF VALVAT. OF VERTE VALVAT. 204 MP AD SLERNAD FILP GTG -L SATAE JOSA VERTET SCHE VALVAR TO VERTE SCHE VALVAR. OF VERTE VALUE	140			AP	SGKEG	PGCVTNTV	LOCATING	AT PROVE	TKPVPDVTA	DAH	KOPV	TA PLAG		LGLICE	KPTGAO			
97 A VEPTGLESN - SE GARENVED DIG JOCKT ET LAVD TO ALEDNEK EN KULSU VA JADOA JAGO KALGEF 104 A VEPTGLESN - SE GARENVED DIG JOCKT ET LAVD TO ALEDNE JSK KESKELSU - VA SVADOT JAGO KALGEF 112 A VEPTGLESN - SE GVEDVE GELGEN SES ALEDNE JNA JAN SE DEL GVA SVAD	94				ΛV	PTGLFSN-S	LCOATNVL	GIGIDCKTS	TIAVDTGAI	DAHOASK	GERPL CC	VA SVAD		OAMO	KAIGTF			
104 A V P T G LF SN = 3L CCARTING P IG UPOKT = TIAVD TGAIS DAR SAKES KPL SG - VA SVAD OT A T CONTACT ON A LG F 112 A V P T G LF SN = 3L CONTACT ON TO T G Z IGONGE A SKA IF DG LS D NA SVAN ZL DPL SG - VA SVAD OF A T CONTACT ON TO T OT CONTACT OT CONTACT ON TO T OT CONTACT ON TO T OT CONT	97				ΛV	PTGLFSN-	LCATNVL	IGVDCKTS	TIAVDTGAI	DAH ASK	SKPLC	VA PVAD-		OALCO	KAIGTF			
112 A VEPTGIYSA - SECONDARCZIGQNERASKA IFDGLS DAN ZAVOLDEL SE-VA ZADOGHTOD IFVGTQ 153 APATSPNGYDA PDGGLIGTD SEL ZAVGZLSGESSISKTPN SAKE DE IZVAS OKAR SETESE ZFTLGAF DXPVGVA 298 LPATSANGYEA PDGGLFGND SES INAVGLSGERATKTPN SAKE DA IZVES OKAR SETESE ELELGAF DXPVGVA 204 MPADSLSRADFILP STG-LSATAE SES AND GLE ZAVE ZVAR SEE ZVA SOERA SE DA IZVES OF ZVA SOERA SEE ZVA SOERA	104				ΛV	PTGLFSN-S	LCCATNVL	LIGYDCKT 5	TIAVDTGAI	DAHOASK	SS KPL CC	VA SVAD-		OTALCO	KAIGTF			
153 AP AT SPNG YD A PDGGL IGT D COSL DYG ZL SGEDS S JSK TPN SAKE DE I DANS OKAR COFUS E JFT	112				ΛV	PTG IY SN - 3	LOCVEDVEC	JIGONCEAS	KSAIFDGLS	ONADAAV	LDPL C	VA SVAD-		OGHLCD	IPVGTQ			
298 LPATSANGYEAPPOGGLEGND ZOSLINDYG ZLSG DORADTK TPNSAKE DA IZNES ZOKAR ZOGISE BLE	153			APATSPI	NGYDA	PDGGLIGT	D SC SL DIAVO	LSGESS	SKTPN SAKE	F DE ICAAS	CO KAR CO	LSEVET-		LGAF	KPVGVTA			
204 MPADSLSRRADFYLPGCTG-LSATAE COSA MODG CLE THOAAVPKTPT SCEHEWA I TAAO DEAR COLTVULGOCUR DTPPGA	298			LPATSAN	NGYEA	PDGGLFGN		LSGDCRAS	TK TPN SAKE	DA INAES	O KAR CO	LSE LE-		LGAF	KPVGVSA			
	204			MPADSL	SRRADFYLP	GTG-LSATA	E CSA STO	LETNOAAV	PKTPTSGEH	VA INA O	O EAR CO	LTVZLG			TPPGA			
205 SPTYGGKYEPOGSA-LYSOAQ COATION OF WALDAR WALSON AND AVAL OF A DWALG	205			SPTY	GGKYEP	GSA-LYSQA	OCATOVE	VACLDCD GV	GARIDNAQH	F VN L CA AK	CORAR CC	AIPVLG		QALLCO	EPEGTN			

Supplementary Figure 6AA. Sequence alignment and corresponding consensus sequence of proteins contained in Region 14 (full sequence)

	1	10	20	30	40	50	60	70	80
Consensus	CXXXXXX	XXPXCCXXB	VIXXXII	зСхх дхххххх х	XXEXXXCAX	XCXXXXCC-	XXPXXGXXLJC	XXXXXXXXAXX	XXXAX
Identity									
84	PGGLLH	IVPLCCATN	VIDVATLI	OCSTRTV PV L TP	AAFOSHCAG	KEKOPVEC-	STRAAGLGLLC	EKPIGTO	
85	CPGGVLHI	VPLCCATN	VIGVATL	C G TRTV PV S T A	A OFOA HCAG	KCKOPVCC-	SLPVAGLGLLC	EKPVGTO	
139	CPGGVTNT	TVPLCCATN	VIGLATLI	GVPTKPVPDV	TAFOAHCAS	K GKO PV CC -	TAPLAGLGLLC	EKPTGAO	
140	CPGCVTNT	TVPLCCATN	VIGLATII	GVPTKPVPDV	TAFOAHCAS	KCKOPVCC-	TAPLAGLGLLC	EKPTGAÕ	
94	CPTGLFS1	-PLCCATN	VLDLIGVI	CKTPTIAVDTG	AIFOAHCAS	KCSKPLCC-	VAPVADQALLC	QKAIGTF	
97	CPTGLFS1	N-PLCCATN	VLDLIGVI	DCKTPTIAVDTG	AIFQAHCAS	KGSKPLCC-	VAPVADQALLC	QKAIGTF	
104	CPTGLFS1	N-PLCCATN	VLDLIGVI	DCKTPTIAVDTG	AIFQAHCAS	KGSKPLCC-	VAPVADQTLLC	QKAIGTF	
112	CPTGIYSI	N-PLCCVTD	VFGVIGQ	CEAPKSAIFDG	LSEQNACAA	VGLDPLCC-	VAPVADQGILC	QIPVGTQ	
98	CPEGLLY	INPLCCDLD	VLGVADVI	CVV PPAKPS SC	KSFGSVCAS	IGRKPRCC-	AVPIAGVALLC	TDPIPAI	
234	CSSC	SNPV CCATD	VLNLANLI	D C AP P ATTPT SI	NAFIDGCAA	G <mark>GQQA</mark> K CC -	LLPILGQALIC	SDVNPTAA	
333	CSSC	ENPV CCATD	VLNLANLI	D C AP P ATTPT SI	NAFIDGCAA	G <mark>GQQA</mark> K CC -	LLPILGQALIC	SDVNPTAA	
72	CSSC	SNPV CCATD	VLNLANLI	DCAPPATTPT SI	NAFIDGCAA	AGQQAKCC-	LIPILGQALIC	SDVNPTADTTA	AKKAL
350	CSSC	TPV CCATD	VLQLANLI	ΟϹΑΡΡΑΙΤΡΤΑΙ	NEFIDTCAE	G <mark>GQQA</mark> K CC -	LIPILGQALIC	S DV N P TA EA PA	TAAAA
188	C A S C	G S PV <mark>CC</mark> A T D	VLNLANLI	DCAPPTTTPADV	NTEIDVCAT	G <mark>GQQA</mark> K CC -	LIPILGQALLC	S DV N P TAAAPA	ΤΡ SA
310	CSSC	GNPV CCATD	VLDLADLI	DCAPPRVTPAGV	NEFIDGCAA	Q <mark>GQQA</mark> K CC -	LIPILGQALIC	S D V N – – – P A P A	A
337	CSSC	GNPV CCATD	VLDLADLI	DCSPPTVTPAGV	NEFIDTCSS	VGQQAKCC-	LIPILGQALIC	S D V N P T A P A P S	AAA
342	CSSC	S P L <mark>C C</mark> A T D	VLNLAILI	DCAAPSVTPAGV	NEFIDTCAG	Q <mark>GQQA</mark> K <mark>CC</mark> -	LIPILGQALLC	SDVNP T APATS	AA
320	CSSC	GNPV CCATD	VLDLADLI	DCAAPSITPTST	DEFINTCAS	AGQQAKCC-	LIPILGQALIC	SDVNPTAPAPS	AA
313	CSSC	GTPL <mark>CC</mark> ATD	VLNVADLI	Y CYNPATTPANV	DELITGCADA	AGQQAKCC-	LIPILGQALIC	NDVSP S AGGAS	тs
90	CTGTNS	SNA Q <mark>CC</mark> A TN	VLGVADL 1	C FN P PTTPT SL	QDEVDI CAA I	DGQQAMCC-	LLPILGQALVC	SEVNP	
387	CSG ADS	SSAV CC ATD		DCAPPPSTPT SI	DDFTSICAS	IGQQARCC-	LLPILEQGLIC	ISPGK_	
352	CSSG	-NPQCCDVD	ATCANT	DCEAPPAAYTDI	KSESDVCAD	VGKINMCC-	DLPVLGQGLIC	SSPDNS	
381	C ASG	-TPQCCDLD	VLGIANL	CQTPPEVPTDI	QSENDICAS	VGQINQCC-	LLPILGQALLC	NSPDGS	
373	CSSAA	INPLCCDVN	VIGVADI	CQVPTIPTSV	PDENDICAS	IGKINMCC-	ILPILGQALLC	ISPDS S T	
30	PALI	DTPLCCQAD	VLGVLDL	TCEARS-DDTSV	SNFEAACAT	TGLTARCC-	TLPLLGEALLC	TTP	
317	PALI	DTELCCQAD	VDGVVDV'	TCEARS-DTSSV	NARKSSCAS	SGTTAKCC-	TLPLAGDALLC	SSP	
174	CPALI	DTPMCCSTR	VLGVADLS	SCSTES-DTSSV	AAENSACVA	TGKTAECC-	TVPVAGQALLC	SSP	
32	CT ALY	TEQCCQAS		ACTPESLSVNSK	QTLVDDCAN'	TGA TAQCC-	VLPIAGQALLC	YDL	
315	CTALY	TEQCCETS	VIGVADL	ACNTESVSVDSE	QSLIND C AK	TGATAECC-	VLPVAGQALLC	YDL	
306	CSSAII	DTAQCCSVS		CVSPHVPPTSV	ESFREICAA	HEQQASCC-	AVPVAGQAVLC	TAASA	
428	CSSALI	DTALCCDVS	VAGVANI	CAAPSTTPTTL	ATEQAICATO	GCOQASCC-	VLPLAGEALLC	PAP	
205	GS-ALYS	SQAQCCATD		DGVGARIDNA	QHEVNLCAA	KEORARCC-	AIPVLGQALLC	QEPEGIN	
153	CPDGGLIC	TEQCOSLD	LVGVLSGI	SSISKTPNSA	KEFQEICAA	SGOKARCC	LSEVFTLGAFC	QKPVGVTA	
298	CPDGGLFC	NEQCOSIN	LVGVLSG	RATKTPNSA	KEFQAICAE	SGOKARCC		QKPVGVSA	
204	GTG-LSI	ATAECCSAN	DGLED	GAAVPKTPTSG	EHEVALOAA	QGQEARCC-	- LTVVLGQGVKC	QTPPGA .	

Supplementary Figure 6AB. Sequence alignment and corresponding consensus sequence of proteins contained in **Region 14** (without N terminus region preceding first cysteine of the eight-cysteine pattern)

	518	514	525	522	533	530	528	529	527	513	523	504	519	511	509	535	515	526	510	501	750 (P	965	447	443	941	454	852	448	944	622	621	1070	702	630	696	626	634	627	625
518	$>\!\!<$	96%	98%	98%	95%	98%	98%	99%	99%	99%	99%	99%	98%	91%	91%	84%	84%	84%	73%	73%	71%	69%	77%	75%	74%	72%	72%	70%	77%	66%	64%	74%	70%	70%	69%	69%	69%	69%	64%
514	96%	$>\!\!<$	95%	95%	93%	95%	95%	96%	96%	96%	96%	96%	95%	89%	89%	81%	81%	81%	71%	71%	70%	66%	75%	74%	73%	70%	70%	69%	74%	65%	63%	71%	69%	69%	68%	68%	68%	68%	63%
525	98%	95%	\geq	100%	98%	99%	99%	98%	98%	98%	99%	99%	98%	90%	90%	83%	83%	83%	73%	73%	71%	69%	75%	74%	73%	70%	70%	69%	75%	66%	64%	74%	70%	70%	70%	70%	70%	70%	65%
522	98%	95%	100%	> <	98%	99%	99%	98%	98%	98%	99%	99%	98%	90%	90%	83%	83%	83%	73%	73%	71%	69%	75%	74%	73%	70%	70%	69%	75%	66%	64%	74%	70%	70%	70%	70%	70%	70%	65%
533	95%	93%	98%	98%	$>\!\!<$	96%	96%	95%	95%	95%	96%	96%	95%	89%	89%	81%	81%	81%	71%	71%	69%	68%	74%	73%	72%	69%	68%	68%	74%	64%	61%	71%	68%	68%	68%	68%	68%	68%	63%
530	98%	95%	99%	99%	96%	\geq	100%	99%	99%	99%	98%	98%	96%	90%	90%	83%	83%	83%	74%	74%	73%	69%	77%	75%	74%	72%	72%	70%	77%	68%	65%	75%	71%	71%	69%	69%	69%	69%	64%
528	98%	95%	99%	99%	96%	100%	\geq	99%	99%	99%	98%	98%	96%	90%	90%	83%	83%	83%	74%	74%	73%	69%	77%	75%	74%	72%	72%	70%	77%	68%	65%	75%	71%	71%	69%	69%	69%	69%	64%
529	99%	96%	98%	98%	95%	99%	99%	\geq	100%	100%	99%	99%	98%	91%	91%	84%	84%	84%	74%	74%	73%	69%	78%	77%	75%	73%	73%	72%	78%	66%	63%	73%	70%	70%	67%	67%	67%	67%	61%
527	99%	96%	98%	98%	95%	99%	99%	100%	> <	100%	99%	99%	98%	91%	91%	84%	84%	84%	74%	74%	73%	69%	78%	77%	75%	73%	73%	72%	78%	66%	63%	73%	70%	70%	67%	67%	67%	67%	61%
513	99%	96%	98%	98%	95%	99%	99%	100%	100%	\geq	99%	99%	98%	91%	91%	84%	84%	84%	74%	74%	73%	69%	78%	77%	75%	73%	73%	72%	78%	68%	65%	75%	71%	71%	69%	69%	69%	69%	64%
523	99%	96%	99%	99%	96%	98%	98%	99%	99%	99%	$>\!\!<$	100%	99%	91%	91%	84%	84%	84%	73%	73%	71%	69%	77%	75%	74%	72%	72%	70%	77%	66%	64%	74%	70%	70%	70%	70%	70%	70%	65%
504	99%	96%	99%	99%	96%	98%	98%	99%	99%	99%	100%	\geq	99%	91%	91%	84%	84%	84%	73%	73%	71%	69%	77%	75%	74%	72%	72%	70%	77%	66%	64%	74%	70%	70%	70%	70%	70%	70%	65%
519	98%	95%	98%	98%	95%	96%	96%	98%	98%	98%	99%	99%	\geq	90%	90%	83%	83%	83%	73%	73%	71%	69%	77%	75%	74%	72%	72%	70%	75%	65%	63%	73%	69%	69%	70%	70%	70%	70%	65%
511	91%	89%	90%	90%	89%	90%	90%	91%	91%	91%	91%	91%	90%	\geq	100%	83%	83%	83%	70%	70%	68%	69%	77%	75%	74%	75%	70%	69%	77%	66%	64%	74%	69%	69%	69%	69%	69%	69%	64%
509	91%	89%	90%	90%	89%	90%	90%	91%	91%	91%	91%	91%	90%	100%	\geq	83%	83%	83%	70%	70%	68%	69%	77%	75%	74%	75%	70%	69%	77%	66%	64%	74%	69%	69%	69%	69%	69%	69%	64%
535	84%	81%	83%	83%	81%	83%	83%	84%	84%	84%	84%	84%	83%	83%	83%	\geq	100%	98%	69%	69%	68%	68%	69%	68%	68%	68%	65%	67%	72%	68%	65%	69%	65%	68%	64%	64%	64%	64%	60%
515	84%	81%	83%	83%	81%	83%	83%	84%	84%	84%	84%	84%	83%	83%	83%	100%	\geq	98%	69%	69%	68%	68%	69%	68%	68%	68%	65%	67%	72%	68%	65%	69%	65%	68%	64%	64%	64%	64%	60%
526	84%	81%	83%	83%	81%	83%	83%	84%	84%	84%	84%	84%	83%	83%	83%	98%	98%	\geq	69%	69%	68%	68%	69%	68%	68%	68%	65%	67%	72%	66%	64%	69%	64%	66%	64%	64%	64%	64%	60%
510	73%	71%	73%	73%	71%	74%	74%	74%	74%	74%	73%	73%	73%	70%	70%	69%	69%	69%	> <	100%	75%	66%	72%	69%	72%	69%	70%	73%	70%	63%	60%	65%	63%	64%	63%	63%	63%	63%	58%
501	73%	71%	73%	73%	71%	74%	74%	74%	74%	74%	73%	73%	73%	70%	70%	69%	69%	69%	100%	\geq	75%	66%	72%	69%	72%	69%	70%	73%	70%	63%	60%	65%	63%	64%	63%	63%	63%	63%	58%
750 (PcaHyd1)	71%	70%	71%	71%	69%	73%	73%	73%	73%	73%	71%	71%	71%	68%	68%	68%	68%	68%	75%	75%	\geq	69%	72%	70%	70%	64%	74%	68%	72%	69%	68%	69%	71%	71%	71%	71%	71%	71%	66%
965	69%	66%	69%	69%	68%	69%	69%	69%	69%	69%	69%	69%	69%	69%	69%	68%	68%	68%	66%	66%	69%	\geq	67%	65%	64%	60%	73%	68%	70%	59%	56%	64%	59%	61%	63%	63%	63%	63%	58%
447	77%	75%	75%	75%	74%	77%	77%	78%	78%	78%	77%	77%	77%	77%	77%	69%	69%	69%	72%	72%	72%	67%	\geq	98%	91%	88%	83%	84%	88%	70%	68%	74%	73%	73%	72%	72%	72%	72%	68%
443	75%	74%	74%	74%	73%	75%	75%	77%	77%	77%	75%	75%	75%	75%	75%	68%	68%	68%	69%	69%	70%	65%	98%	\geq	89%	85%	81%	84%	86%	69%	67%	75%	72%	72%	69%	69%	69%	69%	65%
941	74%	73%	73%	73%	72%	74%	74%	75%	75%	75%	74%	74%	74%	74%	74%	68%	68%	68%	72%	72%	70%	64%	91%	89%	$\geq \leq$	84%	81%	83%	85%	69%	67%	73%	70%	70%	69%	69%	69%	69%	65%
454	72%	70%	70%	70%	69%	72%	72%	73%	73%	73%	72%	72%	72%	75%	75%	68%	68%	68%	69%	69%	64%	60%	88%	85%	84%	\geq	79%	80%	78%	65%	63%	69%	67%	67%	67%	67%	67%	67%	63%
852	72%	70%	70%	70%	68%	72%	72%	73%	73%	73%	72%	72%	72%	70%	70%	65%	65%	65%	70%	70%	74%	73%	83%	81%	81%	79%	\geq	84%	81%	69%	67%	70%	70%	70%	69%	69%	69%	69%	64%
448	70%	69%	69%	69%	68%	70%	70%	72%	72%	72%	70%	70%	70%	69%	69%	67%	67%	67%	73%	73%	68%	68%	84%	84%	83%	80%	84%	\sim	85%	67%	64%	72%	69%	72%	68%	68%	68%	68%	63%
944	77%	74%	75%	75%	74%	77%	77%	78%	78%	78%	77%	77%	75%	77%	77%	72%	72%	72%	70%	70%	72%	70%	88%	86%	85%	78%	81%	85%	\sim	68%	65%	74%	72%	74%	68%	68%	68%	68%	63%
622	66%	65%	66%	66%	64%	68%	68%	66%	66%	68%	66%	66%	65%	66%	66%	68%	68%	66%	63%	63%	69%	59%	70%	69%	69%	65%	69%	67%	68%	~	98%	85%	77%	77%	76%	76%	76%	76%	73%
621	64%	63%	64%	64%	61%	65%	65%	63%	63%	65%	64%	64%	63%	64%	64%	65%	65%	64%	60%	60%	68%	56%	68%	67%	67%	63%	67%	64%	65%	98%	\geq	83%	74%	74%	76%	76%	76%	76%	73%
1070	74%	71%	74%	74%	71%	75%	75%	73%	73%	75%	74%	74%	73%	74%	74%	69%	69%	69%	65%	65%	69%	64%	74%	75%	73%	69%	70%	72%	74%	85%	83%	\geq	79%	78%	80%	79%	79%	78%	73%
702	70%	69%	70%	70%	68%	71%	71%	70%	70%	71%	70%	70%	69%	69%	69%	65%	65%	64%	63%	63%	71%	59%	73%	72%	70%	67%	70%	69%	72%	77%	74%	79%	\sim	96%	85%	84%	84%	83%	78%
630	70%	69%	70%	70%	68%	71%	71%	70%	70%	71%	70%	70%	69%	69%	69%	68%	68%	66%	64%	64%	71%	61%	73%	72%	70%	67%	70%	72%	74%	77%	74%	78%	96%	\sim	84%	84%	85%	84%	78%
696	69%	68%	70%	70%	68%	69%	69%	67%	67%	69%	70%	70%	70%	69%	69%	64%	64%	64%	63%	63%	71%	63%	72%	69%	69%	67%	69%	68%	68%	76%	76%	80%	85%	84%	\geq	99%	99%	98%	92%
626	69%	68%	70%	70%	68%	69%	69%	67%	67%	69%	70%	70%	70%	69%	69%	64%	64%	64%	63%	63%	71%	63%	72%	69%	69%	67%	69%	68%	68%	76%	76%	79%	84%	84%	99%		99%	98%	92%
034	69%	68%	70%	70%	68%	69%	69%	67%	67%	69%	70%	70%	70%	69%	69%	64%	64%	64%	63%	63%	71%	63%	72%	69%	69%	67%	69%	68%	68%	76%	76%	79%	84%	85%	99%	99%		99%	92%
627	69%	68%	70%	70%	68%	69%	69%	67%	67%	69%	70%	70%	70%	69%	69%	64%	64%	64%	63%	63%	71%	63%	72%	69%	69%	67%	69%	68%	68%	76%	76%	78%	83%	84%	98%	98%	99%	\sim	92%
625	64%	63%	65%	65%	63%	64%	64%	61%	61%	64%	65%	65%	65%	64%	64%	60%	60%	60%	58%	58%	66%	58%	68%	65%	65%	63%	64%	63%	63%	73%	73%	73%	78%	78%	92%	92%	92%	92%	

 Table 2A. Identities of region 1 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

 Table 2B. Identities of region 2 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	1037	884	753	751	849	964	438	811	942	842	902	893	930	848	814	639	933	931	899	822	815	951	833	834	821	830	812	442	828
1037	$>\!\!\!<$	56%	57%	48%	56%	51%	54%	55%	54%	57%	57%	56%	54%	52%	51%	57%	54%	62%	60%	58%	56%	53%	59%	53%	54%	58%	54%	64%	62%
884	56%	\geq	51%	45%	56%	50%	54%	56%	58%	59%	58%	53%	57%	56%	53%	59%	52%	58%	56%	49%	49%	49%	54%	46%	47%	47%	51%	56%	60%
753	57%	51%	\geq	68%	67%	55%	56%	57%	51%	57%	48%	52%	48%	57%	57%	52%	54%	58%	51%	54%	51%	51%	63%	54%	63%	57%	57%	63%	60%
751	48%	45%	68%	\geq	64%	55%	59%	60%	51%	56%	49%	54%	46%	54%	56%	56%	52%	54%	46%	54%	49%	47%	57%	49%	52%	49%	49%	56%	53%
849	56%	56%	67%	64%	$>\!\!\!<$	64%	68%	73%	70%	70%	64%	62%	64%	70%	70%	68%	67%	68%	67%	64%	62%	56%	64%	49%	59%	53%	60%	70%	67%
964	51%	50%	55%	55%	64%	\geq	75%	78%	63%	64%	60%	63%	59%	58%	59%	63%	52%	59%	59%	56%	63%	53%	56%	46%	49%	49%	52%	54%	59%
438	54%	54%	56%	59%	68%	75%	$>\!\!<$	75%	63%	68%	56%	60%	60%	54%	63%	64%	52%	58%	56%	53%	62%	57%	60%	51%	54%	49%	52%	63%	65%
811	55%	56%	57%	60%	73%	78%	75%	\geq	69%	69%	59%	63%	64%	67%	69%	69%	58%	62%	64%	60%	58%	57%	60%	51%	54%	52%	56%	63%	69%
942	54%	58%	51%	51%	70%	63%	63%	69%	$\geq <$	73%	59%	58%	62%	62%	59%	64%	57%	62%	63%	57%	57%	46%	56%	47%	51%	51%	51%	59%	63%
842	57%	59%	57%	56%	70%	64%	68%	69%	73%	\geq	54%	52%	59%	57%	57%	60%	58%	64%	60%	63%	58%	51%	53%	48%	51%	48%	54%	62%	64%
902	57%	58%	48%	49%	64%	60%	56%	59%	59%	54%	\geq	68%	65%	67%	65%	68%	58%	64%	59%	57%	46%	51%	51%	41%	46%	46%	49%	53%	53%
893	56%	53%	52%	54%	62%	63%	60%	63%	58%	52%	68%	\geq	65%	60%	64%	62%	52%	59%	59%	52%	53%	47%	52%	48%	53%	49%	47%	56%	53%
930	54%	57%	48%	46%	64%	59%	60%	64%	62%	59%	65%	65%	$>\!$	60%	59%	62%	57%	58%	64%	57%	56%	56%	59%	47%	51%	52%	54%	62%	63%
848	52%	56%	57%	54%	70%	58%	54%	67%	62%	57%	67%	60%	60%	\geq	85%	78%	60%	60%	64%	56%	56%	54%	63%	53%	59%	53%	65%	59%	60%
814	51%	53%	57%	56%	70%	59%	63%	69%	59%	57%	65%	64%	59%	85%	\geq	80%	59%	62%	60%	58%	56%	56%	62%	52%	59%	52%	64%	57%	64%
639	57%	59%	52%	56%	68%	63%	64%	69%	64%	60%	68%	62%	62%	78%	80%	\geq	56%	70%	60%	59%	56%	57%	62%	56%	52%	51%	59%	58%	65%
933	54%	52%	54%	52%	67%	52%	52%	58%	57%	58%	58%	52%	57%	60%	59%	56%	\geq	69%	67%	60%	53%	52%	57%	46%	49%	44%	51%	59%	56%
931	62%	58%	58%	54%	68%	59%	58%	62%	62%	64%	64%	59%	58%	60%	62%	70%	69%	$>\!$	64%	63%	53%	54%	60%	48%	49%	48%	53%	60%	58%
899	60%	56%	51%	46%	67%	59%	56%	64%	63%	60%	59%	59%	64%	64%	60%	60%	67%	64%	\geq	60%	60%	51%	53%	44%	53%	47%	52%	59%	58%
822	58%	49%	54%	54%	64%	56%	53%	60%	57%	63%	57%	52%	57%	56%	58%	59%	60%	63%	60%	$>\!$	54%	57%	54%	51%	57%	57%	56%	60%	58%
815	56%	49%	51%	49%	62%	63%	62%	58%	57%	58%	46%	53%	56%	56%	56%	56%	53%	53%	60%	54%	\geq	51%	56%	47%	49%	48%	56%	58%	62%
951	53%	49%	51%	47%	56%	53%	57%	57%	46%	51%	51%	47%	56%	54%	56%	57%	52%	54%	51%	57%	51%	$>\!\!<$	73%	58%	58%	63%	60%	64%	59%
833	59%	54%	63%	57%	64%	56%	60%	60%	56%	53%	51%	52%	59%	63%	62%	62%	57%	60%	53%	54%	56%	73%	\geq	70%	72%	72%	69%	70%	65%
834	53%	46%	54%	49%	49%	46%	51%	51%	47%	48%	41%	48%	47%	53%	52%	56%	46%	48%	44%	51%	47%	58%	70%	$\geq <$	78%	73%	65%	56%	52%
821	54%	47%	63%	52%	59%	49%	54%	54%	51%	51%	46%	53%	51%	59%	59%	52%	49%	49%	53%	57%	49%	58%	72%	78%	\geq	78%	70%	58%	57%
830	58%	47%	57%	49%	53%	49%	49%	52%	51%	48%	46%	49%	52%	53%	52%	51%	44%	48%	47%	57%	48%	63%	72%	73%	78%	$>\!\!\!<$	69%	62%	57%
812	54%	51%	57%	49%	60%	52%	52%	56%	51%	54%	49%	47%	54%	65%	64%	59%	51%	53%	52%	56%	56%	60%	69%	65%	70%	69%	\geq	67%	59%
442	64%	56%	63%	56%	70%	54%	63%	63%	59%	62%	53%	56%	62%	59%	57%	58%	59%	60%	59%	60%	58%	64%	70%	56%	58%	62%	67%	\geq	70%
828	62%	60%	60%	53%	67%	59%	65%	69%	63%	64%	53%	53%	63%	60%	64%	65%	56%	58%	58%	58%	62%	59%	65%	52%	57%	57%	59%	70%	$>\!\!\!<$

	907	900	910	612	611	610	616	609	436	923	971	970	780	615	453	444	446	926	824	451	641	935	1063	550	911	570	568	877	818	776	809	806	539 (H	860	664	663	1071	1059	1008	1041	1065	817
907	\sim	88%	73%	73%	73%	72%	65%	63%	63%	62%	52%	57%	57%	56%	63%	64%	64%	67%	58%	54%	55%	59%	55%	61%	67%	46%	44%	54%	55%	56%	49%	47%	52%	58%	48%	48%	48%	45%	37%	53%	46%	54%
900	88%	\sim	70%	70%	70%	69%	64%	64%	65%	60%	52%	57%	57%	56%	64%	64%	61%	69%	60%	51%	59%	58%	57%	59%	68%	49%	41%	56%	57%	56%	48%	42%	49%	53%	48%	49%	51%	46%	40%	54%	46%	51%
910	73%	70%	\sim	69%	69%	69%	66%	67%	65%	63%	56%	57%	54%	61%	59%	60%	60%	61%	58%	51%	51%	60%	51%	57%	65%	52%	46%	55%	49%	50%	49%	51%	60%	58%	49%	49%	51%	46%	41%	47%	48%	54%
612	73%	70%	69%	\sim	99%	98%	82%	77%	69%	69%	56%	57%	58%	64%	65%	66%	64%	66%	59%	49%	52%	57%	49%	52%	65%	49%	45%	54%	52%	51%	51%	53%	57%	57%	43%	46%	47%	43%	36%	42%	44%	52%
611	73%	70%	69%	99%	\sim	96%	82%	78%	69%	69%	56%	57%	58%	64%	65%	66%	64%	66%	59%	49%	52%	58%	49%	52%	65%	48%	45%	54%	52%	51%	51%	52%	57%	57%	43%	46%	47%	43%	36%	42%	43%	52%
610	72%	69%	69%	98%	96%	\sim	82%	75%	68%	69%	56%	57%	58%	64%	64%	65%	63%	65%	58%	48%	51%	55%	48%	51%	63%	49%	45%	52%	54%	51%	51%	53%	57%	57%	43%	46%	46%	43%	36%	42%	44%	52%
616	65%	64%	66%	82%	82%	82%		72%	62%	66%	58%	58%	55%	63%	64%	65%	62%	64%	59%	48%	54%	56%	47%	50%	60%	47%	45%	49%	51%	49%	48%	52%	55%	54%	43%	43%	45%	39%	36%	38%	43%	49%
609	63%	64%	67%	77%	78%	75%	72%	\sim	73%	62%	52%	54%	52%	64%	66%	66%	64%	59%	57%	48%	54%	59%	50%	50%	66%	45%	44%	51%	51%	51%	46%	48%	56%	55%	42%	43%	46%	42%	35%	40%	46%	52%
436	63%	65%	65%	69%	69%	68%	62%	73%	\sim	59%	60%	63%	51%	59%	69%	69%	66%	61%	52%	49%	52%	58%	49%	50%	57%	48%	48%	54%	48%	49%	46%	44%	49%	55%	46%	48%	49%	41%	42%	45%	43%	48%
923	62%	60%	63%	69%	69%	69%	66%	62%	59%	\sim	54%	55%	56%	55%	57%	57%	57%	55%	57%	51%	47%	54%	50%	50%	59%	44%	46%	52%	55%	51%	54%	53%	52%	53%	45%	45%	45%	41%	40%	45%	45%	52%
971	52%	52%	56%	56%	56%	56%	58%	52%	60%	54%	5	88%	63%	60%	56%	57%	56%	50%	54%	46%	56%	51%	50%	56%	61%	49%	51%	59%	56%	50%	54%	52%	51%	48%	50%	45%	50%	46%	44%	44%	42%	54%
970	57%	57%	57%	57%	57%	57%	58%	54%	63%	55%	88%		63%	56%	58%	60%	58%	52%	49%	44%	56%	52%	49%	54%	63%	46%	46%	59%	52%	48%	52%	49%	51%	47%	46%	44%	46%	44%	43%	46%	39%	52%
780	57%	57%	54%	58%	5.8%	5.8%	55%	52%	51%	56%	63%	63%	\sim	68%	53%	54%	57%	53%	52%	51%	52%	58%	50%	57%	61%	48%	46%	48%	51%	50%	62%	56%	57%	52%	47%	48%	45%	48%	41%	49%	47%	52%
615	56%	56%	61%	64%	64%	64%	63%	64%	59%	55%	60%	56%	68%		56%	56%	59%	55%	54%	48%	48%	59%	49%	52%	64%	49%	45%	52%	49%	48%	54%	54%	61%	5.8%	43%	45%	44%	45%	37%	46%	48%	53%
453	63%	64%	59%	65%	65%	64%	64%	66%	69%	57%	56%	58%	53%	56%		99%	96%	74%	68%	60%	68%	65%	45%	49%	56%	45%	39%	50%	51%	46%	53%	45%	48%	49%	45%	46%	47%	38%	44%	45%	43%	51%
444	64%	64%	60%	66%	66%	65%	65%	66%	69%	57%	57%	60%	54%	56%	99%		98%	75%	68%	62%	69%	66%	45%	49%	57%	46%	39%	50%	52%	46%	54%	46%	49%	50%	45%	46%	46%	39%	42%	45%	43%	49%
446	64%	61%	60%	64%	64%	63%	62%	64%	66%	57%	56%	5.8%	57%	59%	96%	98%		73%	69%	64%	67%	69%	46%	51%	60%	49%	39%	49%	52%	46%	54%	46%	49%	50%	44%	45%	46%	39%	41%	45%	43%	49%
926	67%	69%	61%	66%	66%	65%	64%	59%	61%	55%	50%	52%	53%	55%	74%	75%	73%	\sim	67%	59%	63%	65%	51%	54%	56%	45%	37%	50%	57%	50%	49%	48%	49%	51%	46%	48%	47%	41%	41%	48%	48%	51%
824	5.8%	60%	5.8%	59%	59%	5.8%	59%	57%	52%	57%	54%	49%	52%	54%	68%	68%	69%	67%		67%	67%	61%	51%	51%	58%	48%	42%	48%	60%	53%	48%	48%	46%	48%	52%	49%	50%	42%	38%	45%	47%	49%
451	54%	51%	51%	49%	49%	48%	48%	48%	49%	51%	46%	44%	51%	48%	60%	62%	64%	59%	67%	\sim	63%	60%	48%	51%	51%	45%	37%	40%	50%	43%	51%	53%	51%	43%	46%	47%	46%	47%	44%	46%	43%	49%
641	55%	59%	51%	52%	52%	51%	54%	54%	52%	47%	56%	56%	52%	48%	68%	69%	67%	63%	67%	63%	\sim	60%	45%	43%	55%	40%	38%	43%	52%	46%	53%	51%	48%	44%	44%	42%	46%	41%	41%	44%	43%	48%
935	59%	5.8%	60%	57%	5.8%	55%	56%	59%	5.8%	54%	51%	52%	5.8%	59%	65%	66%	69%	65%	61%	60%	60%		58%	55%	64%	50%	44%	54%	54%	55%	54%	54%	51%	48%	49%	48%	47%	48%	40%	46%	48%	53%
1063	55%	57%	51%	49%	49%	48%	47%	50%	49%	50%	50%	49%	50%	49%	45%	45%	46%	51%	51%	48%	45%	58%	~	68%	69%	53%	53%	58%	56%	56%	49%	46%	48%	47%	53%	50%	53%	48%	41%	51%	50%	51%
550	61%	59%	57%	52%	52%	51%	50%	50%	50%	50%	56%	54%	57%	52%	49%	49%	51%	54%	51%	51%	43%	55%	68%		67%	56%	\$3%	58%	56%	59%	48%	49%	49%	45%	46%	48%	54%	45%	39%	49%	45%	50%
911	67%	68%	65%	65%	65%	63%	60%	66%	57%	59%	61%	63%	61%	64%	56%	57%	60%	56%	5.8%	51%	55%	64%	69%	67%	~	57%	49%	65%	63%	62%	50%	50%	62%	52%	50%	46%	51%	48%	35%	50%	50%	56%
570	46%	49%	52%	49%	4.8%	49%	47%	45%	4.8%	44%	49%	46%	48%	49%	45%	46%	49%	45%	48%	45%	40%	50%	53%	56%	57%		65%	53%	52%	52%	46%	44%	44%	43%	45%	44%	45%	47%	41%	36%	44%	44%
568	44%	41%	46%	45%	45%	45%	45%	44%	4.8%	46%	51%	46%	46%	45%	39%	39%	39%	37%	42%	37%	3.8%	44%	53%	53%	49%	65%		54%	44%	49%	50%	48%	45%	44%	46%	45%	46%	47%	42%	36%	43%	45%
877	54%	56%	5.5%	54%	54%	52%	49%	51%	54%	52%	59%	59%	48%	52%	50%	50%	49%	50%	48%	40%	43%	54%	58%	58%	65%	53%	54%	3.00	52%	54%	48%	42%	51%	46%	49%	46%	46%	45%	3.8%	45%	48%	49%
818	55%	57%	49%	52%	52%	54%	51%	51%	4.8%	55%	56%	52%	51%	49%	51%	52%	52%	57%	60%	50%	52%	54%	56%	56%	63%	52%	44%	52%	JEN C	64%	46%	50%	46%	45%	48%	48%	47%	43%	45%	48%	43%	49%
776	56%	56%	50%	51%	51%	51%	49%	51%	49%	51%	50%	48%	50%	4.8%	46%	46%	46%	50%	53%	43%	46%	55%	56%	59%	62%	52%	49%	54%	64%		48%	46%	43%	44%	52%	51%	51%	46%	43%	46%	49%	55%
809	49%	48%	49%	51%	51%	51%	48%	46%	46%	54%	54%	52%	62%	54%	53%	54%	54%	49%	48%	51%	53%	54%	49%	48%	50%	46%	50%	48%	46%	48%	<u> </u>	79%	58%	53%	51%	53%	53%	54%	49%	49%	51%	53%
806	47%	42%	51%	53%	52%	53%	52%	48%	44%	53%	52%	40%	56%	54%	45%	46%	46%	48%	48%	53%	51%	54%	46%	49%	50%	44%	48%	42%	50%	46%	79%		65%	55%	48%	48%	48%	46%	45%	44%	4.8%	54%
539 (HGE)	52%	49%	60%	57%	5.7%	57%	55%	56%	49%	52%	51%	51%	57%	61%	48%	40%	40%	40%	46%	51%	48%	51%	48%	49%	62%	44%	45%	51%	46%	43%	5.8%	65%		57%	48%	54%	47%	47%	41%	51%	49%	53%
860	5.8%	53%	5.8%	57%	57%	57%	54%	55%	55%	53%	48%	47%	52%	5.8%	40%	50%	50%	51%	48%	43%	44%	48%	47%	45%	52%	43%	44%	46%	45%	44%	53%	55%	5.7%		49%	55%	53%	43%	41%	48%	51%	49%
664	48%	48%	49%	43%	43%	43%	43%	42%	46%	45%	50%	46%	47%	43%	45%	45%	44%	46%	52%	46%	44%	49%	53%	46%	50%	45%	46%	49%	48%	52%	51%	48%	48%	40%		78%	61%	5.7%	52%	55%	53%	5.8%
663	48%	40%	49%	46%	46%	46%	43%	43%	4.8%	45%	45%	44%	48%	45%	46%	46%	45%	48%	40%	47%	47%	48%	50%	48%	46%	44%	45%	46%	48%	51%	53%	48%	54%	55%	78%		70%	55%	60%	50%	49%	49%
1071	48%	51%	51%	47%	47%	46%	45%	46%	49%	45%	50%	46%	45%	44%	47%	46%	46%	47%	50%	46%	46%	47%	53%	54%	51%	45%	46%	46%	47%	51%	53%	48%	47%	53%	61%	70%		55%	55%	52%	49%	51%
1059	45%	46%	46%	43%	43%	43%	30%	42%	41%	41%	46%	44%	48%	45%	3.8%	30%	30%	41%	47%	47%	41%	48%	48%	45%	48%	47%	47%	45%	43%	46%	54%	46%	47%	43%	57%	55%	55%	33%	57%	5.7%	5.9%	52%
1008	3.7%	40%	41%	36%	36%	36%	36%	35%	42%	40%	44%	43%	41%	37%	44%	47%	41%	41%	38%	44%	41%	40%	41%	30%	35%	41%	47%	38%	45%	43%	40%	45%	41%	41%	5.2%	60%	55%	5.7%		5.2%	47%	46%
1041	53%	54%	47%	42%	42%	42%	3.8%	40%	45%	45%	44%	46%	40%	46%	45%	45%	45%	48%	45%	46%	44%	46%	51%	49%	5.0%	36%	36%	45%	48%	46%	40%	44%	51%	48%	55%	5.0%	52%	57%	5.2%	Jun I	51%	52%
1065	46%	46%	48%	44%	43%	44%	43%	46%	43%	45%	47%	30%	47%	48%	43%	43%	43%	48%	47%	43%	43%	48%	50%	45%	5.0%	44%	43%	48%	43%	40%	51%	48%	40%	51%	53%	40%	40%	50%	47%	5.1%		57%
817	54%	51%	54%	5.2%	52%	52%	40%	5.2%	48%	52%	54%	52%	52%	5394	51%	40%	40%	51%	40%	40%	48%	53%	51%	5.0%	56%	44%	45%	40%	40%	55%	52%	54%	53%	40%	5.8%	40%	51%	52%	46%	52%	5.7%	JIK
																																										-

 Table 2C. Identities of region 3 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

Table 2D. Identities of region 4 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	1067	642	1058	631	650	647	649	648	1056	1034	629	628	635	1054	798	1064	469	460	475	567	885	711	583	915	540	485	484	483	480 (S	541	1060	697	1038
1067	\geq	61%	63%	65%	58%	57%	52%	52%	57%	58%	60%	60%	59%	64%	65%	61%	49%	49%	49%	57%	53%	51%	56%	57%	57%	50%	50%	50%	50%	57%	45%	44%	52%
642	61%	\geq	67%	66%	66%	62%	62%	62%	57%	55%	55%	55%	54%	59%	61%	54%	48%	48%	48%	56%	56%	54%	55%	52%	52%	48%	48%	48%	48%	52%	48%	46%	59%
1058	63%	67%	$>\!\!<$	84%	74%	78%	78%	78%	62%	67%	70%	70%	70%	73%	71%	67%	62%	62%	61%	68%	68%	59%	57%	66%	66%	59%	59%	59%	59%	60%	53%	48%	60%
631	65%	66%	84%	$>\!\!<$	68%	70%	70%	71%	66%	66%	71%	71%	71%	71%	70%	66%	60%	60%	60%	70%	63%	59%	55%	61%	61%	53%	53%	53%	53%	59%	51%	51%	55%
650	58%	66%	74%	68%	\geq	83%	78%	79%	57%	57%	63%	63%	63%	62%	71%	66%	63%	63%	62%	72%	65%	66%	56%	63%	63%	62%	62%	62%	62%	60%	53%	52%	59%
647	57%	62%	78%	70%	83%	$>\!\!<$	78%	79%	57%	60%	60%	60%	60%	63%	70%	62%	65%	65%	63%	70%	66%	66%	56%	66%	66%	60%	60%	60%	60%	64%	54%	51%	59%
649	52%	62%	78%	70%	78%	78%	\geq	99%	56%	54%	59%	59%	59%	59%	66%	62%	68%	68%	67%	63%	73%	64%	59%	63%	63%	56%	56%	56%	56%	58%	52%	48%	57%
648	52%	62%	78%	71%	79%	79%	99%	\geq	57%	55%	60%	60%	60%	59%	67%	63%	70%	70%	68%	65%	73%	65%	59%	63%	63%	56%	56%	56%	56%	58%	51%	47%	57%
1056	57%	57%	62%	66%	57%	57%	56%	57%	$>\!\!<$	89%	74%	74%	74%	82%	78%	57%	60%	60%	60%	60%	59%	63%	54%	54%	54%	52%	52%	52%	52%	57%	53%	52%	55%
1034	58%	55%	67%	66%	57%	60%	54%	55%	89%	\geq	77%	77%	76%	83%	74%	60%	56%	56%	56%	60%	57%	61%	54%	54%	54%	53%	53%	53%	53%	57%	54%	54%	58%
629	60%	55%	70%	71%	63%	60%	59%	60%	74%	77%	\geq	100%	99%	89%	74%	72%	62%	62%	61%	65%	57%	60%	56%	59%	59%	58%	58%	58%	58%	61%	57%	55%	57%
628	60%	55%	70%	71%	63%	60%	59%	60%	74%	77%	100%	\geq	99%	89%	74%	72%	62%	62%	61%	65%	57%	60%	56%	59%	59%	58%	58%	58%	58%	61%	57%	55%	57%
635	59%	54%	70%	71%	63%	60%	59%	60%	74%	76%	99%	99%	$\geq \leq$	88%	74%	71%	62%	62%	61%	63%	57%	60%	55%	59%	59%	57%	57%	57%	57%	61%	57%	55%	57%
1054	64%	59%	73%	71%	62%	63%	59%	59%	82%	83%	89%	89%	88%	\geq	79%	70%	60%	60%	59%	62%	63%	61%	63%	63%	63%	59%	59%	59%	59%	65%	59%	58%	61%
798	65%	61%	71%	70%	71%	70%	66%	67%	78%	74%	74%	74%	74%	79%	\geq	67%	60%	60%	60%	62%	68%	68%	56%	67%	67%	59%	59%	59%	59%	65%	57%	55%	63%
1064	61%	54%	67%	66%	66%	62%	62%	63%	57%	60%	72%	72%	71%	70%	67%	$>\!$	60%	60%	59%	65%	59%	56%	56%	59%	59%	56%	56%	56%	56%	60%	57%	52%	59%
469	49%	48%	62%	60%	63%	65%	68%	70%	60%	56%	62%	62%	62%	60%	60%	60%	$\geq \leq$	100%	99%	67%	60%	63%	57%	63%	63%	62%	62%	62%	62%	60%	57%	53%	52%
460	49%	48%	62%	60%	63%	65%	68%	70%	60%	56%	62%	62%	62%	60%	60%	60%	100%	\geq	99%	67%	60%	63%	57%	63%	63%	62%	62%	62%	62%	60%	57%	53%	52%
475	49%	48%	61%	60%	62%	63%	67%	68%	60%	56%	61%	61%	61%	59%	60%	59%	99%	99%	\geq	67%	60%	63%	57%	63%	63%	60%	60%	60%	60%	59%	57%	53%	52%
567	57%	56%	68%	70%	72%	70%	63%	65%	60%	60%	65%	65%	63%	62%	62%	65%	67%	67%	67%	\geq	55%	63%	59%	65%	65%	67%	67%	67%	67%	59%	55%	53%	57%
885	53%	56%	68%	63%	65%	66%	73%	73%	59%	57%	57%	57%	57%	63%	68%	59%	60%	60%	60%	55%	\geq	75%	61%	64%	64%	57%	57%	57%	57%	61%	55%	51%	54%
711	51%	54%	59%	59%	66%	66%	64%	65%	63%	61%	60%	60%	60%	61%	68%	56%	63%	63%	63%	63%	75%	\geq	63%	62%	62%	65%	65%	65%	65%	62%	56%	52%	58%
583	56%	55%	57%	55%	56%	56%	59%	59%	54%	54%	56%	56%	55%	63%	56%	56%	57%	57%	57%	59%	61%	63%	\geq	62%	62%	56%	56%	56%	56%	58%	51%	46%	54%
915	57%	52%	66%	61%	63%	66%	63%	63%	54%	54%	59%	59%	59%	63%	67%	59%	63%	63%	63%	65%	64%	62%	62%	\geq	100%	70%	70%	70%	70%	72%	55%	52%	65%
540	57%	52%	66%	61%	63%	66%	63%	63%	54%	54%	59%	59%	59%	63%	67%	59%	63%	63%	63%	65%	64%	62%	62%	100%	\geq	70%	70%	70%	70%	72%	55%	52%	65%
485	50%	48%	59%	53%	62%	60%	56%	56%	52%	53%	58%	58%	57%	59%	59%	56%	62%	62%	60%	67%	57%	65%	56%	70%	70%	\geq	100%	100%	100%	61%	54%	50%	60%
484	50%	48%	59%	53%	62%	60%	56%	56%	52%	53%	58%	58%	57%	59%	59%	56%	62%	62%	60%	67%	57%	65%	56%	70%	70%	100%	\geq	100%	100%	61%	54%	50%	60%
483	50%	48%	59%	53%	62%	60%	56%	56%	52%	53%	58%	58%	57%	59%	59%	56%	62%	62%	60%	67%	57%	65%	56%	70%	70%	100%	100%	\geq	100%	61%	54%	50%	60%
480 (SC3)	50%	48%	59%	53%	62%	60%	56%	56%	52%	53%	58%	58%	57%	59%	59%	56%	62%	62%	60%	67%	57%	65%	56%	70%	70%	100%	100%	100%	\geq	61%	54%	50%	60%
541	57%	52%	60%	59%	60%	64%	58%	58%	57%	57%	61%	61%	61%	65%	65%	60%	60%	60%	59%	59%	61%	62%	58%	72%	72%	61%	61%	61%	61%	\geq	57%	55%	61%
1060	45%	48%	53%	51%	53%	54%	52%	51%	53%	54%	57%	57%	57%	59%	57%	57%	57%	57%	57%	55%	55%	56%	51%	55%	55%	54%	54%	54%	54%	57%	\geq	90%	63%
697	44%	46%	48%	51%	52%	51%	48%	47%	52%	54%	55%	55%	55%	58%	55%	52%	53%	53%	53%	53%	51%	52%	46%	52%	52%	50%	50%	50%	50%	55%	90%	\geq	61%
1038	52%	59%	60%	55%	59%	59%	57%	57%	55%	58%	57%	57%	57%	61%	63%	59%	52%	52%	52%	57%	54%	58%	54%	65%	65%	60%	60%	60%	60%	61%	63%	61%	$>\!$

Table 2E. Identities of region 5 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	1068	698	700	1046	1040	1049	1048	1053	1050	1047	1055	1051	1061	589	587	999	575	557	555	918	1066	1000 (477	471	459	476	468	458	586	920	921	710	552	643	916
1068	\geq	89%	79%	84%	85%	87%	82%	72%	77%	73%	52%	54%	54%	46%	50%	49%	52%	48%	51%	53%	66%	48%	48%	49%	49%	49%	54%	51%	56%	54%	48%	48%	50%	63%	51%
698	89%	\geq	80%	84%	83%	84%	77%	72%	78%	78%	56%	57%	57%	49%	52%	49%	52%	48%	51%	57%	68%	53%	53%	54%	54%	54%	57%	52%	59%	55%	52%	52%	52%	59%	57%
700	79%	80%	\geq	74%	73%	79%	70%	72%	75%	74%	49%	51%	50%	44%	44%	45%	46%	45%	44%	53%	62%	46%	46%	48%	48%	49%	51%	49%	54%	48%	45%	48%	46%	55%	45%
1046	84%	84%	74%	$>\!$	99%	88%	77%	77%	76%	73%	60%	60%	57%	54%	54%	51%	51%	47%	56%	53%	67%	50%	50%	51%	51%	50%	54%	49%	56%	54%	53%	54%	52%	60%	52%
1040	85%	83%	73%	99%	\geq	89%	78%	76%	75%	72%	59%	59%	57%	54%	54%	51%	52%	47%	58%	53%	67%	50%	50%	51%	51%	50%	54%	49%	56%	54%	52%	53%	51%	60%	51%
1049	87%	84%	79%	88%	89%	\geq	79%	75%	77%	76%	54%	56%	54%	49%	50%	51%	53%	49%	53%	56%	66%	51%	51%	52%	52%	52%	54%	49%	56%	54%	49%	51%	52%	61%	48%
1048	82%	77%	70%	77%	78%	79%	$>\!\!<$	65%	66%	67%	51%	55%	54%	49%	52%	48%	46%	44%	48%	51%	62%	46%	46%	48%	48%	49%	54%	52%	59%	54%	51%	48%	44%	59%	48%
1053	72%	72%	72%	77%	76%	75%	65%	\geq	89%	82%	59%	57%	55%	53%	55%	48%	48%	45%	51%	51%	61%	45%	45%	46%	46%	43%	51%	47%	55%	51%	49%	49%	51%	61%	53%
1050	77%	78%	75%	76%	75%	77%	66%	89%	\geq	88%	55%	53%	54%	51%	52%	51%	50%	47%	50%	54%	64%	48%	48%	49%	49%	48%	54%	49%	58%	52%	52%	49%	52%	65%	57%
1047	73%	78%	74%	73%	72%	76%	67%	82%	88%	\geq	56%	52%	54%	50%	51%	50%	49%	47%	47%	54%	61%	48%	48%	49%	49%	49%	52%	49%	57%	53%	53%	49%	49%	59%	57%
1055	52%	56%	49%	60%	59%	54%	51%	59%	55%	56%	$\geq \leq$	74%	65%	63%	59%	57%	54%	52%	61%	51%	63%	51%	51%	52%	52%	59%	63%	61%	61%	47%	57%	58%	56%	52%	54%
1051	54%	57%	51%	60%	59%	56%	55%	57%	53%	52%	74%	\geq	68%	63%	60%	61%	60%	55%	66%	52%	61%	52%	52%	54%	54%	55%	59%	54%	55%	47%	55%	52%	51%	48%	55%
1061	54%	57%	50%	57%	57%	54%	54%	55%	54%	54%	65%	68%	\geq	60%	62%	56%	53%	48%	59%	57%	59%	49%	49%	50%	50%	56%	59%	56%	57%	47%	49%	48%	46%	51%	45%
589	46%	49%	44%	54%	54%	49%	49%	53%	51%	50%	63%	63%	60%	\geq	68%	67%	62%	59%	64%	58%	57%	57%	57%	59%	59%	63%	65%	66%	66%	58%	57%	61%	54%	54%	43%
587	50%	52%	44%	54%	54%	50%	52%	55%	52%	51%	59%	60%	62%	68%	\geq	62%	53%	49%	58%	59%	56%	52%	52%	54%	54%	60%	56%	55%	61%	52%	52%	54%	46%	51%	41%
999	49%	49%	45%	51%	51%	51%	48%	48%	51%	50%	57%	61%	56%	67%	62%	\geq	66%	60%	60%	57%	60%	54%	54%	55%	55%	62%	62%	62%	62%	54%	49%	59%	55%	53%	40%
575	52%	52%	46%	51%	52%	53%	46%	48%	50%	49%	54%	60%	53%	62%	53%	66%	\geq	82%	74%	55%	60%	64%	64%	65%	65%	61%	62%	61%	59%	57%	49%	53%	59%	51%	47%
557	48%	48%	45%	47%	47%	49%	44%	45%	47%	47%	52%	55%	48%	59%	49%	60%	82%	\geq	67%	49%	55%	54%	54%	55%	55%	56%	55%	58%	53%	55%	48%	49%	55%	47%	40%
555	51%	51%	44%	56%	58%	53%	48%	51%	50%	47%	61%	66%	59%	64%	58%	60%	74%	67%	\geq	51%	58%	56%	56%	58%	58%	61%	60%	58%	56%	56%	49%	52%	53%	49%	46%
918	53%	57%	53%	53%	53%	56%	51%	51%	54%	54%	51%	52%	57%	58%	59%	57%	55%	49%	51%	\sim	60%	56%	56%	57%	57%	60%	59%	58%	58%	57%	52%	52%	51%	54%	42%
1066	66%	68%	62%	67%	67%	66%	62%	61%	64%	61%	63%	61%	59%	57%	56%	60%	60%	55%	58%	60%	\sim	62%	62%	63%	63%	61%	71%	66%	72%	58%	55%	59%	55%	64%	44%
1000 (VMH2)	48%	53%	46%	50%	50%	51%	46%	45%	48%	48%	51%	52%	49%	57%	52%	54%	64%	54%	56%	56%	62%		100%	99%	99%	82%	73%	74%	67%	61%	51%	54%	60%	53%	48%
4//	48%	53%	46%	50%	50%	51%	46%	45%	48%	48%	51%	52%	49%	57%	52%	54%	64%	54%	56%	56%	62%	100%		99%	99%	82%	73%	74%	67%	61%	51%	54%	60%	53%	48%
4/1	49%	54%	48%	51%	51%	52%	48%	46%	49%	49%	52%	54%	50%	59%	54%	55%	65%	55%	58%	57%	63%	99%	99%	1000	100%	83%	74%	76%	68%	63%	52%	55%	61%	54%	49%
459	49%	54%	48%	51%	51%	52%	48%	46%	49%	49%	52%	54%	50%	59%	54%	55%	65%	55%	58%	5/%	63%	99%	99%	100%	0.2%	85%	74%	76%	68%	63%	52%	55%	61%	54%	49%
470	49%	54%	49%	50%	50%	52%	49%	43%	48%	49%	59%	55%	56%	63%	60%	62%	61%	56%	61%	60%	61%	82%	82%	83%	83%	730/	73%	76%	6/%	64%	53%	57%	60%	51%	44%
408	54%	57%	51%	54%	54%	54%	54%	51%	54%	52%	0.3%	59%	59%	05%	50%	62%	62%	5 000	60%	5 9%	71%	7 3%	7 376	74%	74%	7 3%	0.001	89%	89%	03%	49%	60%	60%	58%	40%
400	51%	52%	49%	49%	49%	49%	52%	47%	49%	49%	61%	54%	50%	66%	55%	62%	61%	58%	58%	58%	7.2%	74%	74% 67M	70%	76%	6 70%	89%	0.41	84%	67%	51%	63%	50%	60%	44%
020	50%	59%	2470	50%	50%	50%	59%	53%	50%	57%	01%	3 3 76	3770	00%	6176	62%	5 976	5370	50%	5 70	7270	07%	07%	00%	60%	67%	69%	6475	6.704	0770	5 376	0376	59%	04%	49%
920	3476	53%	48%	54%	54%	2476	54%	5176	52%	5 376	4770	47%	4/70	5 7%	52%	2476	37%	3 3 76	30%	57%	58%	61% E 1%	01% E 1%	C 3%	5.2%	64% 5.2%	4.0%	04% E 10/	67%	E EW	55%	5 676	33%	30%	49%
710	40/0	52%	4 3 /0	5 4 9/	52%	49% 51%	4 90/	49%	3270	3370	5776	5 30/	49%	6.19/	5270	49/0	49%	40%	49%	52%	50%	51/0	51/0	5276	52%	5 70/	49%	6 10/	6.20/	5 90/	E EW	33%	4270	4 3 /0	49%
552	40%	52%	46%	5.2%	5 1 9/	5 294	40%	49% 51%	49% 52%	49%	56%	51%	46%	5 A 94	34%	59%	50%	49%	52%	51%	59%	60%	60%	61%	61%	60%	60%	61%	5.0%	5.2%	42%	5.7%	37%	5 7%	40%
643	63%	50%	55%	60%	60%	61%	5.0%	61%	65%	5.0%	52%	48%	51%	54%	51%	53%	51%	47%	40%	54%	64%	53%	53%	5.4%	54%	5.1%	5.8%	60%	64%	5.0%	43%	54%	5.7%	37%	46%
016	5 19/	5 9%	35%	E 294	E 10/	4 90/	3 976	5 29/	6 376 E 70/	5 9%	52%	40%	459/	428	4.1%	33%	3 176	4/76	49%	4.29/	04%	3376	3376	34%	3470	3176	38%	4.49	4.0%	4.0%	40%	3476	128	46%	40%
310	3170	37%	** 376	3276	51%	* 070	40%	3376	3776	37%	3476	33%	** 376	* 376	*170	40%	* 7 76	*076	40%	4276	+476	*1876	* 070	+9%	4970	4476	40%	4470	4 976	* 970	4 976	40%	4370	40%	

	990	489	737	988	738	729 (S	728	987	974	986	733	629	659	495	493	670	770	763	768	758	716	715	608	581	876	546	661	657	956
990	\geq	99%	79%	39%	39%	38%	39%	43%	39%	37%	37%	39%	33%	33%	33%	29%	33%	33%	34%	36%	36%	33%	32%	32%	38%	38%	37%	33%	27%
489	99%	\geq	78%	39%	39%	38%	39%	43%	39%	37%	37%	39%	33%	33%	33%	29%	31%	31%	33%	35%	36%	33%	32%	32%	38%	37%	37%	33%	27%
737	79%	78%	$>\!\!\!<$	40%	39%	39%	39%	40%	39%	43%	41%	38%	35%	35%	35%	35%	33%	33%	34%	35%	34%	32%	30%	32%	38%	38%	39%	35%	25%
988	39%	39%	40%	\geq	81%	49%	51%	47%	42%	51%	46%	46%	36%	36%	36%	38%	31%	31%	32%	41%	34%	33%	40%	38%	39%	36%	36%	42%	32%
738	39%	39%	39%	81%	\geq	46%	47%	43%	40%	46%	44%	43%	33%	33%	33%	33%	32%	32%	31%	39%	33%	34%	41%	35%	37%	35%	35%	38%	30%
729 (SC16)	38%	38%	39%	49%	46%	$>\!\!<$	96%	41%	40%	44%	41%	41%	36%	36%	36%	41%	30%	30%	30%	38%	34%	33%	34%	36%	40%	33%	42%	41%	27%
728	39%	39%	39%	51%	47%	96%	\geq	41%	40%	44%	41%	43%	36%	36%	36%	40%	31%	31%	31%	39%	33%	34%	35%	36%	40%	34%	41%	40%	28%
987	43%	43%	40%	47%	43%	41%	41%	\geq	83%	75%	78%	51%	36%	36%	36%	45%	36%	36%	38%	39%	40%	41%	43%	41%	46%	41%	36%	42%	33%
974	39%	39%	39%	42%	40%	40%	40%	83%	\geq	74%	78%	48%	33%	33%	33%	36%	34%	34%	35%	35%	36%	36%	45%	41%	43%	41%	39%	45%	34%
986	37%	37%	43%	51%	46%	44%	44%	75%	74%	\geq	83%	46%	37%	37%	37%	40%	32%	32%	33%	41%	42%	37%	46%	43%	41%	40%	40%	45%	31%
733	37%	37%	41%	46%	44%	41%	41%	78%	78%	83%	$>\!\!\!<$	46%	39%	39%	39%	40%	38%	38%	39%	44%	40%	36%	48%	42%	42%	42%	36%	40%	29%
629	39%	39%	38%	46%	43%	41%	43%	51%	48%	46%	46%	$\geq <$	38%	38%	38%	39%	34%	34%	35%	40%	38%	43%	45%	45%	43%	44%	43%	46%	33%
659	33%	33%	35%	36%	33%	36%	36%	36%	33%	37%	39%	38%	\geq	100%	100%	40%	28%	28%	28%	42%	38%	33%	41%	45%	42%	34%	37%	39%	26%
495	33%	33%	35%	36%	33%	36%	36%	36%	33%	37%	39%	38%	100%	\geq	100%	40%	28%	28%	28%	42%	38%	33%	41%	45%	42%	34%	37%	39%	26%
493	33%	33%	35%	36%	33%	36%	36%	36%	33%	37%	39%	38%	100%	100%	\geq	40%	28%	28%	28%	42%	38%	33%	41%	45%	42%	34%	37%	39%	26%
670	29%	29%	35%	38%	33%	41%	40%	45%	36%	40%	40%	39%	40%	40%	40%	$>\!$	31%	31%	31%	32%	33%	37%	42%	39%	36%	36%	36%	40%	33%
770	33%	31%	33%	31%	32%	30%	31%	36%	34%	32%	38%	34%	28%	28%	28%	31%	\geq	100%	98%	48%	37%	37%	43%	40%	36%	33%	36%	34%	26%
763	33%	31%	33%	31%	32%	30%	31%	36%	34%	32%	38%	34%	28%	28%	28%	31%	100%	\geq	98%	48%	37%	37%	43%	40%	36%	33%	36%	34%	26%
768	34%	33%	34%	32%	31%	30%	31%	38%	35%	33%	39%	35%	28%	28%	28%	31%	98%	98%	\geq	48%	36%	36%	43%	40%	36%	34%	38%	35%	26%
758	36%	35%	35%	41%	39%	38%	39%	39%	35%	41%	44%	40%	42%	42%	42%	32%	48%	48%	48%	\geq	45%	45%	47%	47%	41%	39%	35%	38%	30%
716	36%	36%	34%	34%	33%	34%	33%	40%	36%	42%	40%	38%	38%	38%	38%	33%	37%	37%	36%	45%	\geq	56%	47%	44%	42%	38%	36%	38%	31%
715	33%	33%	32%	33%	34%	33%	34%	41%	36%	37%	36%	43%	33%	33%	33%	37%	37%	37%	36%	45%	56%	\geq	43%	38%	38%	41%	33%	38%	30%
608	32%	32%	30%	40%	41%	34%	35%	43%	45%	46%	48%	45%	41%	41%	41%	42%	43%	43%	43%	47%	47%	43%	\geq	62%	42%	47%	45%	48%	27%
581	32%	32%	32%	38%	35%	36%	36%	41%	41%	43%	42%	45%	45%	45%	45%	39%	40%	40%	40%	47%	44%	38%	62%	\geq	45%	45%	52%	43%	33%
876	38%	38%	38%	39%	37%	40%	40%	46%	43%	41%	42%	43%	42%	42%	42%	36%	36%	36%	36%	41%	42%	38%	42%	45%	\geq	48%	46%	45%	36%
546	38%	37%	38%	36%	35%	33%	34%	41%	41%	40%	42%	44%	34%	34%	34%	36%	33%	33%	34%	39%	38%	41%	47%	45%	48%	\geq	47%	45%	29%
661	37%	37%	39%	36%	35%	42%	41%	36%	39%	40%	36%	43%	37%	37%	37%	36%	36%	36%	38%	35%	36%	33%	45%	52%	46%	47%	\geq	69%	36%
657	33%	33%	35%	42%	38%	41%	40%	42%	45%	45%	40%	46%	39%	39%	39%	40%	34%	34%	35%	38%	38%	38%	48%	43%	45%	45%	69%	\geq	39%
956	27%	27%	25%	32%	30%	27%	28%	33%	34%	31%	29%	33%	26%	26%	26%	33%	26%	26%	26%	30%	31%	30%	27%	33%	36%	29%	36%	39%	$>\!\!\!\!>$

Table 2F. Identities region 6 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	1062	1043	1004	1003	841	839	840	837	423	242	998	983	742	709	675	996	975	966	959	908	892	927
1062	$>\!\!<$	58%	31%	29%	26%	24%	23%	23%	29%	22%	21%	24%	23%	32%	25%	28%	33%	19%	16%	23%	22%	26%
1043	58%	$>\!\!<$	26%	29%	24%	24%	22%	22%	26%	23%	23%	25%	24%	27%	27%	24%	25%	17%	16%	24%	27%	23%
1004	31%	26%	$>\!\!\!<$	59%	25%	22%	21%	27%	24%	25%	23%	22%	22%	23%	22%	28%	27%	17%	18%	20%	20%	19%
1003	29%	29%	59%	$>\!$	24%	23%	22%	34%	21%	27%	20%	25%	25%	23%	23%	24%	27%	20%	20%	22%	21%	20%
841	26%	24%	25%	24%	$>\!\!\!<$	83%	68%	54%	24%	21%	22%	23%	23%	29%	20%	31%	27%	22%	19%	19%	24%	20%
839	24%	24%	22%	23%	83%	\geq	64%	53%	25%	23%	22%	22%	22%	25%	21%	27%	25%	14%	16%	16%	21%	22%
840	23%	22%	21%	22%	68%	64%	\geq	48%	19%	22%	16%	24%	24%	29%	23%	30%	24%	21%	18%	18%	24%	19%
837	23%	22%	27%	34%	54%	53%	48%	$>\!$	20%	23%	21%	24%	24%	27%	24%	29%	26%	17%	17%	16%	22%	23%
423	29%	26%	24%	21%	24%	25%	19%	20%	$>\!\!<$	38%	34%	26%	26%	27%	23%	33%	27%	22%	17%	26%	17%	26%
242	22%	23%	25%	27%	21%	23%	22%	23%	38%	$>\!$	28%	26%	26%	23%	26%	23%	22%	20%	20%	16%	23%	21%
998	21%	23%	23%	20%	22%	22%	16%	21%	34%	28%	$>\!$	22%	22%	27%	20%	22%	25%	15%	20%	24%	16%	21%
983	24%	25%	22%	25%	23%	22%	24%	24%	26%	26%	22%	\geq	99%	34%	36%	24%	25%	21%	21%	28%	28%	20%
742	23%	24%	22%	25%	23%	22%	24%	24%	26%	26%	22%	99%	\geq	33%	35%	24%	25%	21%	21%	28%	27%	20%
709	32%	27%	23%	23%	29%	25%	29%	27%	27%	23%	27%	34%	33%	$\geq\!$	42%	31%	32%	32%	24%	29%	27%	26%
675	25%	27%	22%	23%	20%	21%	23%	24%	23%	26%	20%	36%	35%	42%	$\geq \leq$	23%	26%	21%	19%	25%	26%	29%
996	28%	24%	28%	24%	31%	27%	30%	29%	33%	23%	22%	24%	24%	31%	23%	$\geq\!$	39%	23%	22%	29%	26%	28%
975	33%	25%	27%	27%	27%	25%	24%	26%	27%	22%	25%	25%	25%	32%	26%	39%	\geq	25%	19%	25%	19%	26%
966	19%	17%	17%	20%	22%	14%	21%	17%	22%	20%	15%	21%	21%	32%	21%	23%	25%	$\geq\!$	59%	29%	33%	16%
959	16%	16%	18%	20%	19%	16%	18%	17%	17%	20%	20%	21%	21%	24%	19%	22%	19%	59%	$>\!$	27%	31%	16%
908	23%	24%	20%	22%	19%	16%	18%	16%	26%	16%	24%	28%	28%	29%	25%	29%	25%	29%	27%	$>\!\!\!<$	38%	26%
892	22%	27%	20%	21%	24%	21%	24%	22%	17%	23%	16%	28%	27%	27%	26%	26%	19%	33%	31%	38%	\geq	19%
927	26%	23%	19%	20%	20%	22%	19%	23%	26%	21%	21%	20%	20%	26%	29%	28%	26%	16%	16%	26%	19%	$>\!\!\!<$

Table 2G. Identities region 7 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

Table 2H. Identities region 8 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	425	421	260	24	384	178	937	380	193	88	40	226	38	31	37	36	35	329	286	432	219 (E	389	372	429	803	18	321
425	$>\!$	70%	68%	69%	65%	70%	32%	20%	28%	13%	14%	18%	19%	19%	19%	19%	20%	15%	19%	15%	16%	17%	19%	10%	23%	16%	8%
421	70%	\geq	61%	63%	63%	64%	36%	23%	28%	14%	14%	17%	14%	14%	14%	14%	17%	18%	19%	16%	19%	19%	20%	11%	22%	16%	9%
260	68%	61%	$>\!\!<$	95%	92%	74%	32%	23%	30%	14%	15%	18%	18%	18%	19%	19%	19%	17%	20%	15%	20%	18%	19%	13%	19%	17%	11%
24	69%	63%	95%	\geq	92%	72%	33%	24%	29%	14%	15%	18%	18%	18%	19%	19%	19%	17%	20%	15%	20%	18%	19%	12%	19%	17%	11%
384	65%	63%	92%	92%	\geq	73%	34%	23%	31%	13%	14%	16%	18%	18%	19%	19%	19%	17%	20%	15%	20%	18%	19%	12%	20%	17%	9%
178	70%	64%	74%	72%	73%	$>\!$	33%	21%	28%	13%	14%	19%	16%	16%	18%	18%	18%	17%	22%	16%	20%	16%	19%	14%	22%	19%	7%
937	32%	36%	32%	33%	34%	33%	$>\!\!<$	24%	28%	15%	15%	13%	20%	20%	17%	17%	20%	20%	23%	21%	14%	19%	13%	15%	22%	19%	11%
380	20%	23%	23%	24%	23%	21%	24%	$>\!$	34%	28%	23%	23%	20%	20%	18%	18%	20%	13%	18%	14%	26%	15%	19%	18%	18%	22%	12%
193	28%	28%	30%	29%	31%	28%	28%	34%	$>\!\!<$	18%	26%	20%	17%	17%	20%	20%	17%	14%	16%	13%	25%	14%	16%	19%	23%	23%	9%
88	13%	14%	14%	14%	13%	13%	15%	28%	18%	$>\!\!\!<$	55%	47%	16%	18%	18%	19%	16%	14%	14%	15%	24%	10%	14%	17%	20%	27%	9%
40	14%	14%	15%	15%	14%	14%	15%	23%	26%	55%	$>\!$	45%	12%	14%	14%	14%	15%	15%	14%	17%	23%	11%	16%	13%	16%	25%	8%
226	18%	17%	18%	18%	16%	19%	13%	23%	20%	47%	45%	\geq	17%	18%	17%	17%	15%	15%	14%	18%	23%	13%	12%	18%	17%	27%	8%
38	19%	14%	18%	18%	18%	16%	20%	20%	17%	16%	12%	17%	$>\!\!\!<$	97%	77%	79%	77%	14%	13%	20%	17%	15%	12%	24%	17%	21%	7%
31	19%	14%	18%	18%	18%	16%	20%	20%	17%	18%	14%	18%	97%	\geq	80%	81%	80%	14%	13%	20%	18%	15%	12%	23%	18%	21%	7%
37	19%	14%	19%	19%	19%	18%	17%	18%	20%	18%	14%	17%	77%	80%	\geq	99%	73%	12%	12%	19%	20%	14%	14%	23%	18%	18%	7%
36	19%	14%	19%	19%	19%	18%	17%	18%	20%	19%	14%	17%	79%	81%	99%	\geq	73%	12%	12%	19%	20%	14%	14%	23%	18%	17%	7%
35	20%	17%	19%	19%	19%	18%	20%	20%	17%	16%	15%	15%	77%	80%	73%	73%	$>\!$	14%	13%	19%	18%	15%	12%	25%	18%	20%	6%
329	15%	18%	17%	17%	17%	17%	20%	13%	14%	14%	15%	15%	14%	14%	12%	12%	14%	\geq	56%	45%	14%	28%	27%	11%	16%	13%	11%
286	19%	19%	20%	20%	20%	22%	23%	18%	16%	14%	14%	14%	13%	13%	12%	12%	13%	56%	\geq	42%	14%	29%	33%	12%	16%	14%	8%
432	15%	16%	15%	15%	15%	16%	21%	14%	13%	15%	17%	18%	20%	20%	19%	19%	19%	45%	42%	$>\!$	16%	25%	22%	12%	14%	17%	8%
219 (EAS)	16%	19%	20%	20%	20%	20%	14%	26%	25%	24%	23%	23%	17%	18%	20%	20%	18%	14%	14%	16%	$>\!\!\!<$	17%	19%	20%	17%	20%	10%
389	17%	19%	18%	18%	18%	16%	19%	15%	14%	10%	11%	13%	15%	15%	14%	14%	15%	28%	29%	25%	17%	\geq	26%	13%	16%	16%	12%
372	19%	20%	19%	19%	19%	19%	13%	19%	16%	14%	16%	12%	12%	12%	14%	14%	12%	27%	33%	22%	19%	26%	$>\!$	9%	8%	11%	11%
429	10%	11%	13%	12%	12%	14%	15%	18%	19%	17%	13%	18%	24%	23%	23%	23%	25%	11%	12%	12%	20%	13%	9%	$\geq \leq$	17%	19%	8%
803	23%	22%	19%	19%	20%	22%	22%	18%	23%	20%	16%	17%	17%	18%	18%	18%	18%	16%	16%	14%	17%	16%	8%	17%	$>\!$	20%	9%
18	16%	16%	17%	17%	17%	19%	19%	22%	23%	27%	25%	27%	21%	21%	18%	17%	20%	13%	14%	17%	20%	16%	11%	19%	20%	$>\!$	7%
321	8%	9%	11%	11%	9%	7%	11%	12%	9%	9%	8%	8%	7%	7%	7%	7%	6%	11%	8%	8%	10%	12%	11%	8%	9%	7%	$>\!\!<$

Table 2I. Identities region 9 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	866	861	854	853	424	592	358	229	177 (D	165	262	261	21	356	224	377	183 (M	225	973	380	294	330	207	34	427	304
866	$>\!\!\!<$	83%	76%	39%	26%	35%	18%	19%	14%	14%	11%	12%	12%	16%	16%	13%	17%	16%	25%	24%	20%	16%	16%	14%	12%	19%
861	83%	\geq	73%	39%	25%	30%	16%	18%	13%	13%	11%	13%	13%	16%	16%	11%	15%	18%	25%	25%	19%	19%	17%	16%	13%	20%
854	76%	73%	\geq	39%	26%	31%	18%	18%	12%	12%	11%	12%	12%	17%	18%	11%	17%	15%	25%	21%	20%	19%	19%	17%	15%	19%
853	39%	39%	39%	$>\!\!\!<$	36%	28%	16%	16%	12%	12%	12%	11%	11%	13%	13%	10%	13%	15%	32%	20%	15%	17%	17%	17%	13%	15%
424	26%	25%	26%	36%	\geq	27%	15%	16%	17%	17%	14%	15%	15%	15%	15%	9%	13%	12%	27%	21%	16%	17%	18%	14%	13%	14%
592	35%	30%	31%	28%	27%	$>\!\!\!<$	15%	15%	18%	18%	16%	14%	14%	19%	18%	10%	15%	15%	31%	22%	19%	17%	16%	15%	11%	18%
358	18%	16%	18%	16%	15%	15%	$>\!\!<$	94%	41%	41%	45%	31%	31%	19%	19%	13%	16%	12%	15%	13%	14%	12%	10%	10%	15%	8%
229	19%	18%	18%	16%	16%	15%	94%	\geq	43%	43%	47%	31%	31%	19%	19%	13%	16%	12%	15%	15%	15%	12%	11%	10%	14%	9%
177 (DewA)	14%	13%	12%	12%	17%	18%	41%	43%	\geq	100%	60%	39%	39%	22%	20%	13%	15%	12%	13%	16%	10%	17%	11%	11%	13%	10%
165	14%	13%	12%	12%	17%	18%	41%	43%	100%	\geq	60%	39%	39%	22%	20%	13%	15%	12%	13%	16%	10%	17%	11%	11%	13%	10%
262	11%	11%	11%	12%	14%	16%	45%	47%	60%	60%	$>\!\!<$	49%	49%	16%	16%	14%	17%	13%	15%	13%	11%	16%	14%	10%	9%	11%
261	12%	13%	12%	11%	15%	14%	31%	31%	39%	39%	49%	\geq	99%	16%	18%	12%	15%	10%	11%	10%	8%	12%	14%	9%	11%	9%
21	12%	13%	12%	11%	15%	14%	31%	31%	39%	39%	49%	99%	\geq	16%	18%	12%	15%	10%	11%	10%	8%	12%	14%	9%	11%	9%
356	16%	16%	17%	13%	15%	19%	19%	19%	22%	22%	16%	16%	16%	\geq	92%	12%	15%	13%	14%	17%	18%	16%	10%	14%	9%	13%
224	16%	16%	18%	13%	15%	18%	19%	19%	20%	20%	16%	18%	18%	92%	\geq	12%	15%	13%	14%	15%	18%	15%	11%	14%	9%	13%
377	13%	11%	11%	10%	9%	10%	13%	13%	13%	13%	14%	12%	12%	12%	12%	\geq	69%	17%	8%	11%	12%	14%	10%	10%	5%	8%
183 (MPG1)	17%	15%	17%	13%	13%	15%	16%	16%	15%	15%	17%	15%	15%	15%	15%	69%	\geq	26%	11%	16%	15%	17%	13%	12%	7%	11%
225	16%	18%	15%	15%	12%	15%	12%	12%	12%	12%	13%	10%	10%	13%	13%	17%	26%	$>\!$	14%	15%	13%	13%	12%	15%	15%	12%
973	25%	25%	25%	32%	27%	31%	15%	15%	13%	13%	15%	11%	11%	14%	14%	8%	11%	14%	\geq	26%	22%	19%	14%	15%	14%	16%
380	24%	25%	21%	20%	21%	22%	13%	15%	16%	16%	13%	10%	10%	17%	15%	11%	16%	15%	26%	$>\!$	18%	21%	14%	17%	13%	13%
294	20%	19%	20%	15%	16%	19%	14%	15%	10%	10%	11%	8%	8%	18%	18%	12%	15%	13%	22%	18%	$>\!$	13%	17%	12%	12%	16%
330	16%	19%	19%	17%	17%	17%	12%	12%	17%	17%	16%	12%	12%	16%	15%	14%	17%	13%	19%	21%	13%	$>\!\!\!<$	17%	13%	11%	15%
207	16%	17%	19%	17%	18%	16%	10%	11%	11%	11%	14%	14%	14%	10%	11%	10%	13%	12%	14%	14%	17%	17%	$>\!\!<$	11%	11%	13%
34	14%	16%	17%	17%	14%	15%	10%	10%	11%	11%	10%	9%	9%	14%	14%	10%	12%	15%	15%	17%	12%	13%	11%	\geq	31%	19%
427	12%	13%	15%	13%	13%	11%	15%	14%	13%	13%	9%	11%	11%	9%	9%	5%	7%	15%	14%	13%	12%	11%	11%	31%	$>\!\!<$	14%
304	19%	20%	19%	15%	14%	18%	8%	9%	10%	10%	11%	9%	9%	13%	13%	8%	11%	12%	16%	13%	16%	15%	13%	19%	14%	$>\!\!\!<$

	407	155	201	383	23	255	159	160	163	164	162	161	185	175	209	272	71	276
407	>	99%	98%	32%	34%	35%	35%	35%	35%	35%	34%	35%	38%	38%	40%	37%	34%	33%
155	99%	$>\!$	99%	31%	33%	35%	35%	35%	35%	35%	34%	35%	38%	38%	40%	37%	34%	33%
201	98%	99%	$>\!\!<$	32%	34%	35%	35%	35%	35%	35%	34%	35%	39%	39%	41%	37%	34%	33%
383	32%	31%	32%	$>\!\!<$	78%	35%	35%	35%	34%	35%	33%	34%	34%	35%	37%	25%	26%	27%
23	34%	33%	34%	78%	$>\!$	37%	37%	37%	36%	36%	35%	36%	35%	35%	36%	27%	27%	25%
255	35%	35%	35%	35%	37%	$>\!\!<$	99%	99%	99%	99%	96%	96%	53%	53%	45%	33%	33%	38%
159	35%	35%	35%	35%	37%	99%	$>\!$	98%	98%	98%	95%	95%	52%	52%	44%	33%	33%	38%
160	35%	35%	35%	35%	37%	99%	98%	$>\!\!\!<$	98%	98%	95%	95%	52%	52%	45%	33%	33%	38%
163	35%	35%	35%	34%	36%	99%	98%	98%	$>\!$	98%	95%	95%	53%	53%	45%	33%	33%	38%
164	35%	35%	35%	35%	36%	99%	98%	98%	98%	\geq	95%	95%	52%	52%	46%	34%	34%	39%
162	34%	34%	34%	33%	35%	96%	95%	95%	95%	95%	\geq	91%	51%	51%	44%	33%	33%	39%
161	35%	35%	35%	34%	36%	96%	95%	95%	95%	95%	91%	\geq	53%	53%	46%	33%	33%	38%
185	38%	38%	39%	34%	35%	53%	52%	52%	53%	52%	51%	53%	$>\!\!<$	99%	66%	34%	34%	34%
175	38%	38%	39%	35%	35%	53%	52%	52%	53%	52%	51%	53%	99%	$>\!$	66%	34%	34%	34%
209	40%	40%	41%	37%	36%	45%	44%	45%	45%	46%	44%	46%	66%	66%	$>\!\!\!<$	33%	33%	34%
272	37%	37%	37%	25%	27%	33%	33%	33%	33%	34%	33%	33%	34%	34%	33%	$>\!\!<$	91%	76%
71	34%	34%	34%	26%	27%	33%	33%	33%	33%	34%	33%	33%	34%	34%	33%	91%	\geq	76%
276	33%	33%	33%	27%	25%	38%	38%	38%	38%	39%	39%	38%	34%	34%	34%	76%	76%	\sim

Table 2J. Identities region 10 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	241	240	379	266	343	361	230	213	217	191	194	180	422
241	$>\!\!\!<$	82%	41%	42%	44%	48%	44%	41%	39%	37%	40%	43%	47%
240	82%	$>\!\!<$	44%	42%	43%	50%	48%	46%	42%	39%	40%	51%	49%
379	41%	44%	$>\!\!<$	96%	50%	52%	54%	46%	44%	48%	44%	49%	47%
266	42%	42%	96%	$>\!\!\!<$	50%	50%	51%	45%	45%	46%	43%	50%	47%
343	44%	43%	50%	50%	$>\!\!<$	52%	54%	42%	39%	44%	39%	45%	50%
361	48%	50%	52%	50%	52%	$>\!\!<$	87%	50%	48%	49%	49%	49%	52%
230	44%	48%	54%	51%	54%	87%	$>\!\!<$	54%	46%	50%	49%	51%	55%
213	41%	46%	46%	45%	42%	50%	54%	$>\!\!<$	46%	43%	38%	49%	46%
217	39%	42%	44%	45%	39%	48%	46%	46%	$>\!\!<$	55%	56%	54%	55%
191	37%	39%	48%	46%	44%	49%	50%	43%	55%	$>\!\!<$	51%	45%	46%
194	40%	40%	44%	43%	39%	49%	49%	38%	56%	51%	$>\!\!<$	52%	51%
180	43%	51%	49%	50%	45%	49%	51%	49%	54%	45%	52%	$>\!\!<$	56%
422	47%	49%	47%	47%	50%	52%	55%	46%	55%	46%	51%	56%	$>\!\!<$

 Table 2K. Identities region 11 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)

	26	19	385	25	265	20	346	323	189 (RodA)	248	195	216	176
26	$>\!\!<$	99%	97%	87%	90%	89%	90%	79%	81%	82%	79%	82%	75%
19	99%	$>\!\!<\!\!<$	98%	86%	90%	90%	91%	80%	82%	83%	80%	83%	76%
385	97%	98%	$>\!$	88%	89%	88%	90%	80%	82%	83%	80%	83%	75%
25	87%	86%	88%	$>\!\!\!<$	81%	80%	80%	74%	76%	77%	77%	78%	72%
265	90%	90%	89%	81%	$>\!\!<$	99%	90%	80%	82%	84%	82%	80%	79%
20	89%	90%	88%	80%	99%	$>\!\!\!<\!\!\!<$	91%	80%	82%	84%	82%	80%	80%
346	90%	91%	90%	80%	90%	91%	$>\!\!<$	82%	84%	86%	83%	82%	79%
323	79%	80%	80%	74%	80%	80%	82%	$>\!\!<$	98%	96%	92%	86%	74%
189 (RodA)	81%	82%	82%	76%	82%	82%	84%	98%	$>\!\!\!<$	98%	94%	88%	76%
248	82%	83%	83%	77%	84%	84%	86%	96%	98%	\geq	96%	89%	77%
195	79%	80%	80%	77%	82%	82%	83%	92%	94%	96%	\geq	87%	76%
216	82%	83%	83%	78%	80%	80%	82%	86%	88%	89%	87%	$>\!\!<$	72%
176	75%	76%	75%	72%	79%	80%	79%	74%	76%	77%	76%	72%	$>\!\!<$

 Table 2L. Identities region 12 (without N-terminus region preceding first cysteine of the eight-cysteine pattern)