Supporting Information

Isolation and characterization of novel peptides from *Momordica cochinchinensis* seeds

Lai Y. Chan, Conan K. L. Wang, Jodie M. Major, Kathryn P. Greenwood, Richard J. Lewis, David J. Craik, and Norelle L. Daly*

The University of Queensland, Institute for Molecular Bioscience, Brisbane QLD 4072, Australia.

* To whom correspondence should be addressed:

The University of Queensland,

Institute for Molecular Bioscience,

Brisbane, QLD 4072, Australia.

Tel: +61-7-3346 2021; fax: +61-7-3346 2101

Email address: n.daly@uq.edu.au

Table of contents

page

Table 1S. Sequence fragments from the enzymatic digestion of peptides from *M. cochinchinensis.*

peptide	peptide sequenced	enzyme used[a]	relative abundance (%)[b]	theoretical mass (Da)	experimental mass (Da)	Δ m (Da)[c]
MCoCC-1 (100[d])	S12 - S33	T	5	2112.89	2113.00	0.11
	C19 - S33	T	49	1492.68	1492.80	0.12
	G1 - R11	T	6	1196.55	1196.60	0.05
	R11 - S33	C	57	2269.00	2268.90	0.10
	G1 - S33	C	64	3291.43	3291.30	0.13
MCoCC-2 (100[d])	G1 - R11	T	7	1165.55	1165.60	0.05
	C19 -S32	T	56	1405.65	1405.60	0.05
	G1 - F10	C	39	1009.45	1009.42	0.03
	R11 - S32	C	100	2181.96	2181.87	0.09

[a]Enzymes used included trypsin (T) and chymotrypsin (C).
[b]The abundance relative to the most intense ion recorded in the TOF scan in the range 400-2000 Da
[c]The absolute difference between the theoretical and experimental mass
[d]Percentage (%) sequence coverage

Table 2S. The comparison of amino acid composition for MCoCC-1 and MCoCC-2 [1]

AMINO ACID	% OF MOLE		RESIDUES/MOL			
			Experimental		Theoretical	
	MCoCC-1	MCoCC-2	MCoCC-1	MCoCC-2	MCoCC-1	MCoCC-2
Histidine (H)	-	0.97	-	0.15	-	0
Serine (S)	12.83	8.48	3.03	1.91	4	3
Arginine (R)	7.73	6.35	1.10	0.86	2	2
Glycine (G)[a]	30.34	28.86	10.04	9.09	8	8
Aspartic acid (D)	1.33	3.33	0.25	0.59	0	0
Glutamic acid (E)/Glutamine (Q)[b]	7.89	4.48	1.33	0.72	2	1
Threonine (T)	7.11	6.53	1.48	1.29	2	2
Alanine (A)	0.80	2.84	0.22	0.75	0	0
Proline (P)	7.72	9.82	1.67	2.02	2	3
Lysine (K)	4.07	3.99	0.69	0.65	1	1
Tyrosine (Y)	-	2.44	-	0.32	-	0
Methionine (M)	-	2.15	-	0.34	-	0
Valine (V)	7.60	7.70	1.61	1.55	2	2
Isoleucine (I)[c]	3.88	3.63	0.73	0.65	1	1
Leucine (L)[c]	4.90	3.74	0.93	0.67	1	1
Phenylalanine (F)	3.81	4.69	0.57	0.67	1	1
Cysteine (C)[d]	ND	ND	ND	ND	6	6
Tryptophan (W)[d]	ND	ND	ND	ND	1	1

[a]Glycine is often overestimated in amino acid analysis
[b]Glutamic acid and glutamine cannot be differentiated one from the other
[c]Leucine and isoleucine (isobaric residues) are yet to be differentiated
[d]ND: Not analyzed by amino acid analysis

[1]This research has been facilitated by access to the Australian Proteome Analysis Facility established under the Australian Governments Major National Research Facilities program.

Figure 3S. 2D NMR spectra of MCoCC-1. The amide region of the TOCSY (top) and fingerprint region of the NOESY (bottom) spectrum are shown in the spectra below. The sample was prepared in 90% H_2O/10% D_2O and measured at 290 K. The spin systems are shown in the TOCSY spectrum while H_α-H_N connectivities are shown in the NOESY. **N.B.** Several TOCSY peaks (C2, S12, C13 and T22) are not be visible at this contour level but are observed at lower levels.

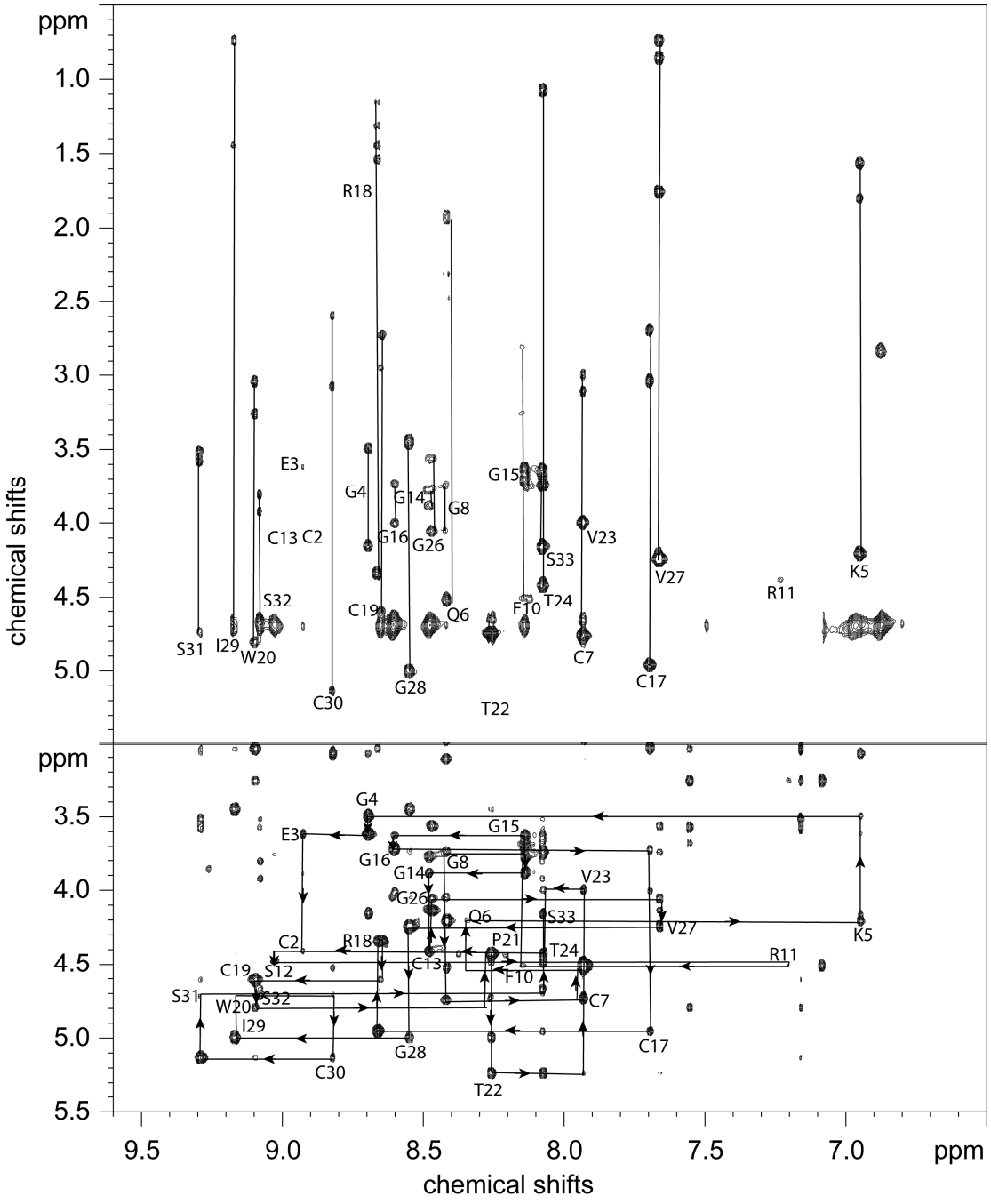

Table 4S. Chemical shifts of MCoCC-1 (ppm) at 290K

Residue	NH	Hα	Hβ	Others
C2	8.927	4.414	3.037, 2.805	-
E3	8.927	3.624	1.928, 1.868	Hγ 2.315, 2.193
G4	8.695	3.500, 4.163	-	-
K5	6.948	4.213	1.802, 1.569	Hγ 1.343, 1.278; γCH_2 1.538
Q6	8.415	4.531	-	Hγ 2.480, 2.318; NH_2 7.494, 6.797; γCH_2 1.929
C7	7.932	4.753	3.117, 2.999	-
G8	8.418	3.744, 4.053	-	-
L9	-	3.684	1.180, 0.9470	Hγ 1.024; δCH_3 0.6205, 0.5583
F10	8.146	4.514	3.260, 2.810	7.120, 7.084, 7.203
R11	7.228	4.387	-	-
S12	-	4.483	3.76	-
C13	9.027	4.417	2.747, 2.544	-
G14	8.478	3.890, 3.778	-	-
G15	8.142	3.629, 3.710	-	-
G16	8.6	3.729, 4.008	-	Hδ 4.089, 3.600
C17	7.696	4.964	3.044, 2.695	-
R18	8.661	4.345	1.538, 1.446	Hγ 1.312, 1.156; γCH_2 2.823
C19	8.651	4.609	2.952, 2.720	-
W20	9.096	4.802	3.263, 3.049	Hδ 7.160; Hε 10.03, 7.555
P21	-	4.43	2.206, 1.834	δCH_2 3.687, 3.574; γCH_2 1.909
T22	8.258	5.244	4.492	γCH_3 1.166
V23	7.931	4.001	2.2	γCH_3 0.8817
T24	8.075	4.425	3.746	γCH_3 1.078
P25	-	4.132	2.141, 1.770	Hδ 4.089, 3.600; Hγ 2.008, 1.858
G26	8.47	3.565, 4.059	-	-
V27	7.661	4.251	1.758	γCH_3 0.7368, 0.8548
G28	8.549	3.451, 5.003	-	-
I29	9.17	4.731	1.444	γCH_2 1.211, 0.9730; γCH_3 0.7380
C30	8.821	5.137	3.082, 2.595	-
S31	9.291	4.731	3.582, 3.519	-
S32	9.078	4.662	3.926, 3.809	-
S33	8.076	4.163	3.666	-

Table 5S. Structural statistics of the MCoCC-1 structures calculated with Set 3 disulfide connectivity

MCoCC-1	
Pairwise RMSD[a] (Å)	
Mean global backbone	0.90 ± 0.25
Mean global heavy chain	1.53 ± 0.50
Mean global backbone (residue 16-33)	0.40 ± 0.15
Mean global heavy chain (residues 16-33)	0.76 ± 0.22
Experimental data	
Distance restraints	195
Long range	69
Medium range	36
Sequential	90
H-bonding restraints	8
Dihedral restraints	13
NOE violations > 0.3 Å	0
Dihedral violations > 3.0°	0
Energies ($kJmol^{-1}$)	
Overall	-1186.4 ± 14.25
Bonds	6.21 ± 0.64
Angles	25.7 ± 2.25
Improper	5.55 ± 9.56
van der Waals	-93.6 ± 7.83
NOE	12.0 ± 3.22
cDIH	0.69 ± 0.28
Dihedral	158.06 ± 5.16
Electrostatic	-1300.95 ± 13.51
RMSD	
Bonds (Å)	3.75E-03 ± 1.92E-04
Angles (degrees)	4.60E-01 ± 2.01E-02
Improper (degrees)	3.96E-01 ± 3.37E-02
NOE	3.40E-02 ± 4.48E-03
cDIH	6.62E-01 ± 1.36E-01
Ramachandran (%)	
Most favoured	80.4
Additionally allowed	19.6
Generously allowed	0
Disallowed	0

[a] RMSD, root mean square deviation