

# Supporting Information

## Exploring a structural data mining approach to design linkers for head-to-tail peptide cyclization

Yasaman Karami,<sup>†</sup> Samuel Murail,<sup>†</sup> Julien Giribaldi,<sup>‡</sup> Benjamin Lefranc,<sup>¶</sup> Florian Defontaine,<sup>§</sup> Olivier Lesouhaitier,<sup>§</sup> Jérôme Leprince,<sup>¶</sup> Sjoerd de Vries,<sup>\*,†</sup> and Pierre Tufféry<sup>\*,†</sup>

<sup>†</sup>*Université Paris Cité, CNRS UMR 8251, INSERM ERL U1133, 75013 Paris, France*

<sup>‡</sup>*Institut des Biomolécules Max Mousseron, UMR 5247, Université de Montpellier-CNRS, 34293 Montpellier, France.*

<sup>¶</sup>*Université de Rouen Normandie, INSERM U1239 NorDiC, Neuroendocrine, Endocrine and Germinal Differentiation and Communication, INSERM US51 HeRacLeS, F-76000 Rouen, France.*

<sup>§</sup>*Université de Rouen Normandie, UR CBSA, Research Unit Bacterial Communication and Anti-infectious Strategies, 27000 Evreux, France.*

E-mail: sjoerd.de-vries@inserm.fr; pierre.tuffery@u-paris.fr

# Database search

We previously introduced the BCLoopSearch protocol, to mine large protein structure datasets and retrieve loop candidates, given two disjoint fragments (loop flanks).<sup>1</sup> It is based on a Binet-Cauchy (BC) kernel and a Rigidity score:

$$BC(X, Y) = \frac{\det(X^T Y)}{\sqrt{\det(X^T X) \det(Y^T Y)}} \quad (\text{S1})$$

where  $X$  and  $Y$  are  $C_\alpha$  coordinates of the flanks and dataset fragments, respectively and they are centered at the origin. Note that a BC score of 1 indicates a perfect match. *Rigidity* score  $R(X, Y)$  is defined as:

$$R'(X, Y) = \max_{1 \leq i \leq N} \|X_i - Y_i\| \quad (\text{S2})$$

$$R(X, Y) = \max\{R'(X, Y), \|\|X_N - X_1\| - \|Y_N - Y_1\|\|\} \quad (\text{S3})$$

where  $X_i$  and  $Y_i$  are  $C_\alpha$  coordinates of the  $i$ th residues of the flanks and dataset fragments and  $\|\cdot\|$  is the euclidean norm. Rigidity score is the maximum variation of intra-distances between: *(i)* residues and geometric center and *(ii)* intra-distances between terminal  $C_\alpha$ . In addition, we also measured the RMSD between query and candidate flanks for the fragments returned. In total, four cutoffs values related to *(i)* flank size, *(ii)* flank BC score, *(iii)* flank Rigidity and *(iv)* flank RMSD, have been considered to limit the number of loop candidates. In this study we used: a flank size of 4 residues, Rigidity  $\leq 2.5$ , flank RMSD  $\leq 4 \text{ \AA}$  and the minimal flank BC score cutoff of 0.8.

## CyBase benchmark

CyBase (<http://www.cybase.org.au/>)<sup>2,3</sup> provides a set of existing naturally occurring cyclic peptides. Presently, 64 3D structures of cyclic peptides from 25 different species are reported. We applied a filtering step on the list to keep only those that are *i*) head-to-tail cyclized, *ii*) without modified amino acids and *iii*) not identical (filtering out entries with identical sequences), resulting in a final set of 35 structures. Residues from the N- and/or C-terminal extremities of each cyclic peptide were removed to generate linear peptides (here by N- and C-terminal extremity, we refer to the head and tail residues from the sequence). We considered all possible combinations of truncating two to seven residues from the N- and/or C-termini (*i.e.*, removing two residues from N-terminus or two residues from C-terminus or one residue from each side), generating 33 different linear peptides from every cyclic target. We also excluded the cases where the size of generated linear peptide was less than 8 residues, that is size limit of our protocol. Finally we obtained a total of 1147 linear peptides, where the corresponding linkers are in the range of 2-7 residue long. The details of those structures are reported in **Supplementary Table 1**.

We applied both our protocol and Rosetta NGK to the CyBase test set to model all the linkers. Over the 1147 cases, both our data-mining and Rosetta NGK failed to model the linker for 6 different cases (0.5%). In fact, our protocol identified candidates in all cases, but discarded all the candidates with a correct geometry but a non satisfactory sequence similarity in 6 cases. Thus overall, in terms of ability to identify linkers, the data-mining strategy seems to perform as well as a pure *ab initio* procedure. Then, we compared both protocols using the 1135 over 1147 (99%) cases for which the linker could be modelled by both methods. All heavy backbone atoms (N, C, C $\alpha$ , O) were considered. The local RMSD corresponds to RMSD obtained by superposing the model linker on the native conformation using a best fit procedure, whereas the global RMSD corresponds to RMSD observed after superposing the linear part of the peptide (*i.e.*, without the linker). The best RMSD over the top 20 predictions by each method were retained.

Table 1: **The list of cyclic structures from CyBase.** Structures with identical sequences were discarded and only one representative was considered. For each cyclic peptide, we generated a total of 33 linear peptides by truncating two to seven residues from N- and/or C-term. The total number of linear peptides for each target, as well as those modelled with PEP-Cyclizer and Rosetta NGK are reported.

Protein name	Class	Type	PDBcode	size	#linkers	#linkers (PEP-Cyclizer)
kalata-B1	Cyclotide	NMR	1NB1	29	33	33
kalata-B1	Cyclotide	NMR	1K48	29	33	33
kalata-B1	Cyclotide	NMR	1KAL	29	33	33
[P20D,V21K]-kalata-B1	Cyclotide	NMR	2F2I	29	33	33
[W19K,-P20N,-V21K]-kalata-B1	Cyclotide	NMR	2F2J	29	33	33
kalata-B2	Cyclotide	NMR	1PT4	29	33	33
kalata-B5	Cyclotide	NMR	2KUX	30	33	32
kalata-B7	Cyclotide	NMR	2JWM	29	33	33
kalata-B7	Cyclotide	NMR	2M9O	29	33	33
kalata-B8	Cyclotide	NMR	2B38	31	33	33
kalata-B12	Cyclotide	NMR	2KVX	28	33	32
cycloviolacin-O1	Cyclotide	NMR	1NBJ	30	33	32
cycloviolacin-O1	Cyclotide	NMR	1DF6	30	33	33
cycloviolacin-O2	Cyclotide	NMR	2KNM	30	33	32
cycloviolacin-O14	Cyclotide	NMR	2GJ0	31	33	33
MCoTI-II	Squash-trypsin-inhibitor	XRAY	4GUX	34	33	33
circulin-A	Cyclotide	NMR	1BH4	30	33	32
circulin-B	Cyclotide	NMR	2ERI	31	33	33
kB1[GHFRWG;23-28]	Cyclotide	NMR	2LUR	29	33	32
[Ala1,15]kB1	Cyclotide	NMR	1N1U	29	33	33
des(24-28)kB1	Cyclotide	NMR	1ORX	24	33	33
SFTI-1	BBI-like-trypsin-inhibitor	XRAY	3P8F	14	25	25
Ent-AS-48	Bacterial	XRAY	1O82	70	33	33
vhl-1	Cyclotide	NMR	1ZA8	31	33	33
vhl-2	Cyclotide	NMR	2KUK	30	33	33
varv-peptide-F	Cyclotide	NMR	2K7G	29	33	33
varv-peptide-F	Cyclotide	XRAY	3E4H	29	33	33
BiKK	BBI-like-trypsin-inhibitor	NMR	2BEY	16	33	33
RTD-1	Primate	NMR	1HVZ	18	33	33
palicourein	Cyclotide	NMR	1R1F	37	33	33
vhr1	Cyclotide	NMR	1VB8	30	33	31
tricyclon-A	Cyclotide	NMR	1YP8	33	33	33
Cter-M	Cyclotide	NMR	2LAM	29	33	33
MCo-PMI	Squash-trypsin-inhibitor	NMR	2M86	51	33	33
Carnocyclin-A	Bacterial	NMR	2KJF	60	33	33
total					1147	1141

Table 2: **The RMSD values for all the linkers of each structure from CyBase.** The average local and global RMSD values are measured over the backbone atoms (N, C, C $\alpha$ , O) for the linkers modelled by both PEP-Cyclizer and Rosetta NGK. For each cyclic peptide, we generated a total of 33 linear peptides by truncating two to seven residues from N- and/or C-terminal extremities. For each target, the number of linear peptides that were cyclized by both PEP-Cyclizer and Rosetta NGK are reported (out of the total 33 linkers).

Protein name	number of linkers	local RMSD (Å)		global RMSD (Å)	
		PEP-Cyclizer	NGK	PEP-Cyclizer	NGK
kalata-B1	33	0.53±0.22	0.56±0.37	1.17±0.45	1.05±0.37
kalata-B1	32	0.69±0.34	0.70±0.57	2.05±1.40	1.77±0.57
kalata-B1	32	0.71±0.26	1.04±0.55	1.45±0.47	1.71±0.47
[P20D,V21K]-kalata-B1	33	0.65±0.37	0.52±0.38	1.40±0.70	0.99±0.38
[W19K,-P20N,-V21K]-kalata-B1	33	0.54±0.28	0.71±0.40	1.31±0.70	1.24±0.40
kalata-B2	33	0.49±0.17	0.40±0.38	1.12±0.47	0.73±0.38
kalata-B5	32	0.74±0.45	0.45±0.57	1.83±1.53	0.88±0.57
kalata-B7	33	0.58±0.19	0.82±0.65	1.15±0.32	1.13±0.65
kalata-B7	33	0.79±0.36	0.56±0.58	2.11±1.52	1.45±0.58
kalata-B8	33	1.12±0.57	1.14±0.61	2.57±1.13	2.11±0.61
kalata-B12	32	0.74±0.38	0.69±0.37	1.57±1.07	1.40±0.37
cycloviolacin-O1	32	1.08±0.68	0.43±0.17	2.83±1.98	0.91±0.17
cycloviolacin-O1	33	0.98±0.43	1.15±0.43	2.14±1.24	1.99±0.43
cycloviolacin-O2	32	0.75±0.43	0.38±0.39	2.07±1.76	0.75±0.39
cycloviolacin-O14	33	0.94±0.57	0.79±0.89	2.61±1.91	2.01±0.89
MCoTI-II	33	0.73±0.43	0.31±0.44	1.70±0.78	0.64±0.44
circulin-A	33	0.97±0.34	0.88±0.34	2.37±0.96	1.59±0.34
circulin-B	33	1.02±0.44	0.37±0.17	2.22±0.91	0.61±0.17
kB1[GHFRWG;23-28]	32	1.08±0.45	1.16±0.40	2.61±1.23	2.18±0.40
[Ala1,15]kB1	33	0.93±0.39	0.88±0.33	1.94±0.73	1.22±0.33
des(24-28)kB1	32	1.34±0.55	1.60±0.64	2.44±0.81	2.88±0.64
SFTI-1	25	0.50±0.41	0.22±0.14	1.48±1.19	0.57±0.14
Ent-AS-48	33	0.54±0.29	0.14±0.08	1.16±0.53	0.23±0.08
vhl-1	33	0.87±0.46	0.36±0.27	1.88±1.01	0.66±0.27
vhl-2	33	0.53±0.25	0.71±0.74	1.24±0.52	1.11±0.74
varv-peptide-F	33	0.61±0.42	0.51±0.64	1.86±1.62	1.24±0.64
varv-peptide-F	33	0.50±0.21	0.41±0.28	1.17±0.52	0.79±0.28
BiKK	33	0.64±0.47	1.05±0.75	1.64±1.27	1.87±0.75
RTD-1	33	0.72±0.37	0.74±0.43	1.98±0.84	1.77±0.43
palicourein	33	1.12±0.36	1.11±0.32	2.24±1.02	1.93±0.32
vhr1	31	0.97±0.58	0.61±0.34	2.36±1.59	1.23±0.34
tricyclon-A	33	0.79±0.30	0.68±0.48	1.64±0.57	1.13±0.48
Cter-M	33	0.60±0.42	0.54±0.69	1.62±1.43	1.45±0.69
MCo-PMI	31	1.11±0.48	1.09±0.51	2.33±1.22	1.93±0.51
Carnocyclin-A	33	0.52±0.26	0.23±0.13	1.21±0.47	0.48±0.13

Table 3: **RMSD and ranks over the CyBase test set.** For every case the best RMSD out of top 20 and the top 1 were considered. The average and standard deviations of best local ( $lRMSD_{20}$ ,  $lRMSD_1$ ) and global ( $gRMSD_{20}$ ,  $gRMSD_1$ ) RMSD values are reported for every gap size.

linker size (# gaps)		2 (101)	3 (139)	4 (175)	5(208)	6 (241)	7 (271)
$lRMSD_{20}$	PEP-Cyclizer	0.32±0.19	0.51±0.22	0.66±0.28	0.77±0.35	0.92±0.44	1.11±0.49
	Rosetta NGK	0.26±0.33	0.38±0.36	0.52±0.43	0.66±0.48	0.83±0.54	1.04±0.66
$gRMSD_{20}$	PEP-Cyclizer	1.38±0.57	1.45±0.50	1.52±0.59	1.64±0.78	2.02±1.25	2.55±1.70
	Rosetta NGK	0.73±0.78	0.84±0.70	0.97±0.68	1.15±0.81	1.60±1.51	1.93±1.74
$lRMSD_1$	PEP-Cyclizer	0.64±0.29	0.94±0.39	1.14±0.50	1.21±0.57	1.43±0.67	1.79±0.80
	Rosetta NGK	0.34±0.35	0.55±0.53	0.73±0.61	0.89±0.63	1.12±0.72	1.35±0.80
$gRMSD_1$	PEP-Cyclizer	2.67±1.40	2.80±1.47	3.06±1.85	3.06±2.18	3.99±2.67	4.90±3.14
	Rosetta NGK	0.89±0.89	1.14±0.99	1.44±1.47	1.68±1.74	2.18±1.87	2.65±2.25

Table 4: **Summary of cyclic linkers for conotoxins.** Data is collected from<sup>4</sup> and additional details are added from the mentioned references. The last column reports the pdb code of the available engineered cyclic peptides. linkers sequences in bold correspond to the functional variants that were considered in this study.

name	linear peptide	linker	activity	structure	stability	p
$\alpha$ -Conotoxin MII	1m2c (1mii)	GGAAG (cMII-5) <sup>5</sup> <b>GGAAGG (cMII-6)</b> <sup>5</sup> GAGGAAG (cMII-7) <sup>5</sup>	not active similar similar	not similar similar similar	- improved improved	
$\alpha$ -Conotoxin ImI	1cnl	A <sup>6</sup> $\beta$ A <sup>6</sup> AG <sup>6</sup> AGG <sup>6</sup>	- - - -	- - - -	slightly improved improved slightly improved slightly improved	
$\alpha$ -Conotoxin Vc1.1	2h8s	GGAAG <sup>7</sup> <b>GGAAGG</b> <sup>7</sup>	substantial loss similar/higher	similar similar	- improved	
$\alpha$ -Conotoxin RgIA	2JUT	GAA <sup>8</sup> GAAG <sup>8</sup> GAAGG <sup>8</sup> <b>GGAAGG</b> <sup>8</sup> <b>GGAAGAG</b> <sup>8</sup>	reduced reduced reduced similar similar	not similar not similar similar similar similar	- - - improved improved	
$\alpha$ -Conotoxin AuIB	1mxn 1m xp	A <sup>9</sup> AG <sup>9</sup> AGG <sup>9</sup> <b>AGGG</b> <sup>9</sup> GGAAG <sup>9</sup> GAGAAG <sup>9</sup> GGAGGAG <sup>9</sup> <b>GGAA</b> <sup>10</sup> AGAGA <sup>10</sup> GGAAGG <sup>10</sup> GGAAAGG <sup>10</sup>	reduced reduced reduced reduced reduced reduced reduced reduced reduced reduced	- - - - - - - similar similar similar -	- improved improved improved improved improved improved improved improved improved improved	
$\chi$ -Conotoxin MrIA	2ew4	<b>AG</b> <sup>11</sup> <b>RGD</b> <sup>12</sup>	similar similar	similar similar	improved improved	
$\omega$ -Conotoxin MVIIA	1mvi	<b>GGPG</b> <sup>13</sup>	-	-	-	
Conotoxin gm9a	1ixt	<b>GLP</b> <sup>14</sup>	-	similar	similar	
Conotoxin bru9a	-	<b>GLP</b> <sup>14</sup>	-	-	similar	

Table 5: Average ranks of the cyclic linkers for conotoxins, using forward-backtrack algorithm.

name	linear peptide	linker	ranks
$\alpha$ -Conotoxin MII	1m2c (1mii)	GGAAG (cMII-5) <b>GGAAGG (cMII-6)</b> <b>GAGGAAG (cMII-7)</b>	19/32 40/64 8/128
$\alpha$ -Conotoxin ImI	1cnl	A $\beta$ A AG AGG	- - 2/4 3/8
$\alpha$ -Conotoxin Vc1.1	2h8s	GGAAG <b>GGAAGG</b>	23/32 33/64
$\alpha$ -Conotoxin RgIA	2JUT	GAA GAAG GAAGG <b>GGAAGG</b> <b>GGAAGAG</b>	8/8 14/16 28/32 24/64 21/128
$\alpha$ -Conotoxin AuIB	1mxn 1mxp	A <b>AG</b> AGG <b>AGGG</b> GGAAG GAGAAG GGAGGAG <b>GGAA</b> AGAGA GGAAGG GGAAAGG	- 1/4 2/8 6/16 25/32 24/64 8/128 14/16 3/32 37/64 35/128
$\chi$ -Conotoxin MrIA	2ew4	<b>AG</b> <b>RGD</b>	2/4 7/27
$\omega$ -Conotoxin MVIIA	1mvi	<b>GGPG</b>	3/16
Conotoxin gm9a	1ixt	<b>GLP</b>	8/27
Conotoxin bru9a	-	<b>GLP</b>	-



Table 6: The RMSD between the 7 UII models generated by MD (M1-M7) and 5 UII models generated by PEP-FOLD (M1-M5), used as input to PEP-Cyclizer.

<b>MD</b>		<b>M1</b>												
		<b>M2</b>	<b>M2</b>											
		<b>M3</b>	1.10	<b>M3</b>										
		<b>M4</b>	3.20	3.34	<b>M4</b>									
		<b>M5</b>	2.64	2.91	2.59	<b>M5</b>								
		<b>M6</b>	2.23	2.52	2.25	1.45	<b>M6</b>							
		<b>M7</b>	2.49	2.70	2.43	1.38	1.00	1.15	<b>M7</b>					
		<b>M8</b>	2.53	2.81	2.50	1.70	1.22	1.26	0.83	<b>M8</b>				
PEP-FOLD	<b>M1</b>	1.98	2.34	2.03	1.96	1.36	1.16	1.45	1.69	<b>M1</b>				
	<b>M2</b>	1.92	2.18	2.16	3.42	2.85	2.53	2.93	2.86	<b>M2</b>				
	<b>M3</b>	1.79	2.16	1.97	2.98	2.27	2.00	2.49	2.38	1.87	<b>M3</b>			
	<b>M4</b>	2.32	2.52	2.64	3.44	3.11	2.78	2.65	2.73	2.79	2.67	2.99	<b>M4</b>	
	<b>M5</b>	2.07	2.51	2.49	4.12	3.44	3.09	3.35	3.36	2.97	2.12	2.30	<b>M5</b>	

Table 7: Peptide characterization

Compound	Code	Sequence	Formula	$MH^+$ calc. <sup>a</sup>	$MH^+$
hU-II	LV-4001	ETPDCFWKYCV	$C_{64}H_{85}N_{13}O_{18}S_2$	1388.56	1388
cyclo-[GAG]hUII	LV-4130	(GAGETPDCFWKYCV)	$C_{71}H_{94}N_{16}O_{20}S_2$	1555.63	1555
cyclo-[AGG]hUII	LV-4131	(AGGETPDCFWKYCV)	$C_{71}H_{94}N_{16}O_{20}S_2$	1555.63	1555
cyclo-[GGA]hUII	LV-4132	(GGAETPDCFWKYCV)	$C_{71}H_{94}N_{16}O_{20}S_2$	1555.63	1555
cyclo-[GPA]hUII	LV-4133	(GPAETPDCFWKYCV)	$C_{74}H_{98}N_{16}O_{20}S_2$	1595.66	1595
Nrf2(76-85)	LV-5554	LDEETGEFL	$C_{46}H_{69}N_9O_{19}$	1052.47	1052
Cyclo-[PAA]Nrf2(76-85)	LV-5562	(PAALDEETGEFL)	$C_{57}H_{84}N_{12}O_{21}$	1273.59	1273
Cyclo-[AGG]Nrf2(76-85)	LV-5565	(AGGLDEETGEFL)	$C_{53}H_{78}N_{12}O_{21}$	1219.54	1219

<sup>a</sup>Calculated  $MH^+$ . <sup>b</sup>Observed  $MH^+$  measured by MALDI-ToF mass spectrometry.

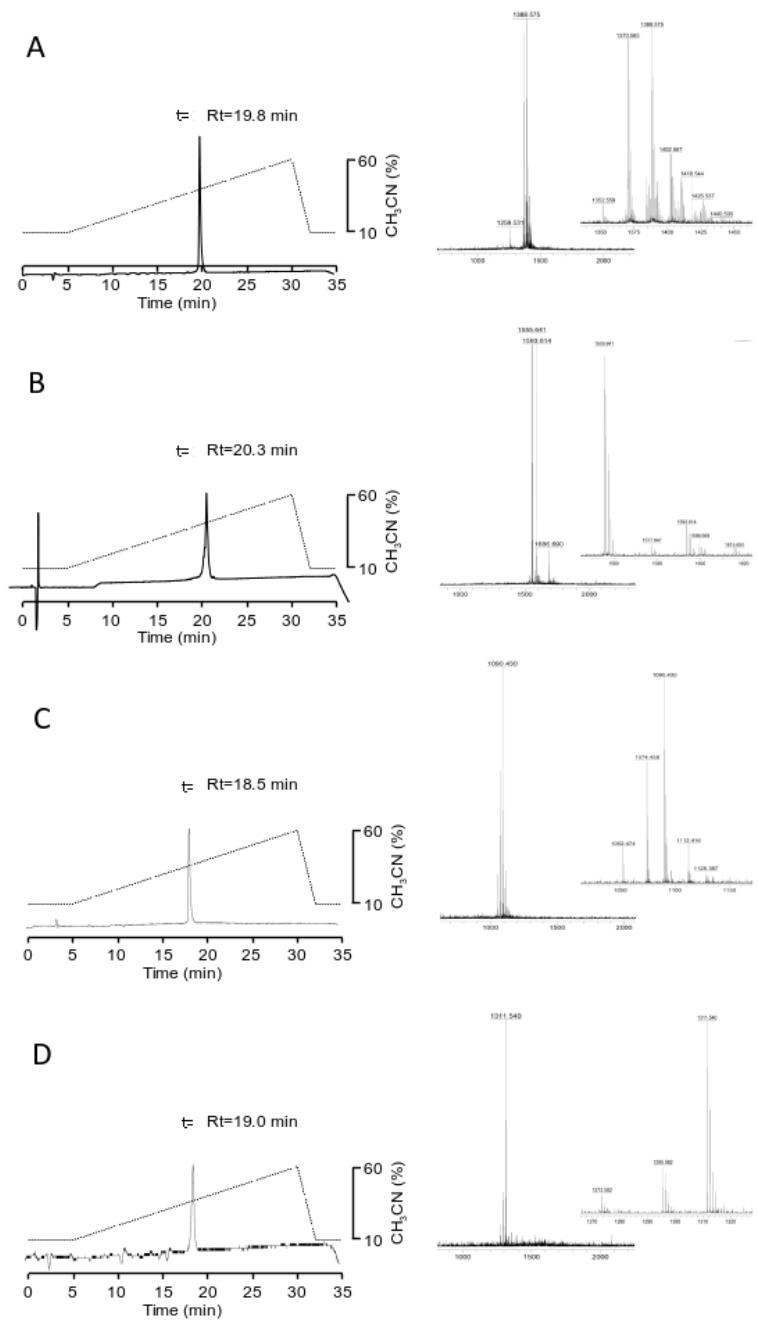


Figure 1: MALDI-Tof MS and RP-HPLC analysis of synthesized peptides. A: UII - LV-4001, B: cyclo[AGG]UII - LV-4131, C: Nrf2(76-85) - LV-5554, D: cyclo[PAA]Nrf2(76-85)-LV-5562.

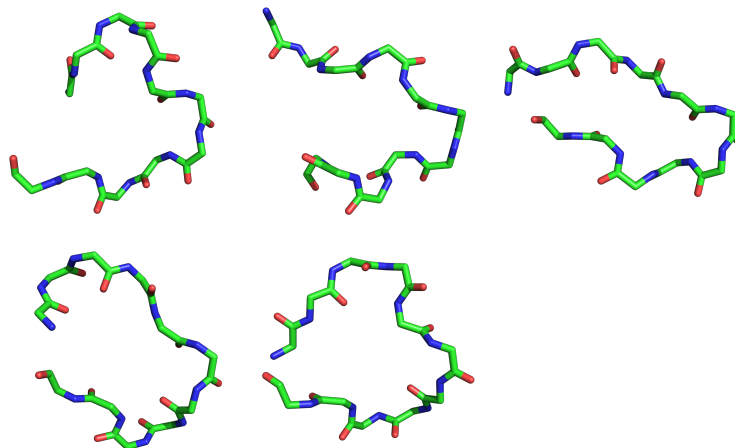


Figure 2: **UII models generated by PEP-FOLD.** The set of 5 UII models generated by PEP-FOLD and refined by MD simulations are shown here.

## References

- (1) Guyon, F.; Martz, F.; Vavrusa, M.; Bécot, J.; Rey, J.; Tufféry, P. BCSearch: fast structural fragment mining over large collections of protein structures. *Nucleic Acids Res* **2015**, *43*, W378–382.
- (2) Mulvenna, J. P.; Wang, C.; Craik, D. J. CyBase: a database of cyclic protein sequence and structure. *Nucleic Acids Res* **2006**, *34*, D192–194.
- (3) Wang, C. K.; Kaas, Q.; Chiche, L.; Craik, D. J. CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res* **2008**, *36*, D206–210.
- (4) Cemazar, M.; Kwon, S.; Mahatmanto, T.; Ravipati, A. S.; Craik, D. J. Discovery and applications of disulfide-rich cyclic peptides. *Curr Top Med Chem* **2012**, *12*, 1534–1545.
- (5) Clark, R. J.; Fischer, H.; Dempster, L.; Daly, N. L.; Rosengren, K. J.; Nevin, S. T.; Meunier, F. A.; Adams, D. J.; Craik, D. J. Engineering stable peptide toxins by means of backbone cyclization: stabilization of the alpha-conotoxin MII. *Proc Natl Acad Sci U S A* **2005**, *102*, 13767–13772.
- (6) Armishaw, C. J.; Dutton, J. L.; Craik, D. J.; Alewood, P. F. Establishing regiocontrol of disulfide bond isomers of alpha-conotoxin ImI via the synthesis of N-to-C cyclic analogs. *Biopolymers* **2010**, *94*, 307–313.

- (7) Clark, R. J.; Jensen, J.; Nevin, S. T.; Callaghan, B. P.; Adams, D. J.; Craik, D. J. The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew Chem Int Ed Engl* **2010**, *49*, 6545–6548.
- (8) Halai, R.; Callaghan, B.; Daly, N. L.; Clark, R. J.; Adams, D. J.; Craik, D. J. *J Med Chem* **2011**, *54*, 6984–6992.
- (9) Armishaw, C. J.; Jensen, A. A.; Balle, L. D.; Scott, K. C.; Sérensen, L.; Strømgaard, K. Improving the stability of  $\alpha$ -conotoxin AuIB through N-to-C cyclization: the effect of linker length on stability and activity at nicotinic acetylcholine receptors. *Antioxid. Redox Signal.* **2011**, *14*, 65–76.
- (10) Lovelace, E. S.; Gunasekera, S.; Alvarmo, C.; Clark, R. J.; Nevin, S. T.; Grishin, A. A.; Adams, D. J.; Craik, D. J.; Daly, N. L. Stabilization of  $\alpha$ -conotoxin AuIB: influences of disulfide connectivity and backbone cyclization. *Antioxid. Redox Signal.* **2011**, *14*, 87–95.
- (11) Lovelace, E. S.; Armishaw, C. J.; Colgrave, M. L.; Wahlstrom, M. E.; Alewood, P. F.; Daly, N. L.; Craik, D. J. Cyclic MrIA: a stable and potent cyclic conotoxin with a novel topological fold that targets the norepinephrine transporter. *J. Med. Chem.* **2006**, *49*, 6561–6568.
- (12) Dekan, Z.; Wang, C. I.; Andrews, R. K.; Lewis, R. J.; Alewood, P. F. Conotoxin engineering: dual pharmacophoric noradrenaline transport inhibitor/integrin binding peptide with improved stability. *Org. Biomol. Chem.* **2012**, *10*, 5791–5794.
- (13) Hemu, X.; Taichi, M.; Qiu, Y.; Liu, D. X.; Tam, J. P. Biomimetic synthesis of cyclic peptides using novel thioester surrogates. *Biopolymers* **2013**, *100*, 492–501.
- (14) Akcan, M.; Clark, R. J.; Daly, N. L.; Conibear, A. C.; de Faoite, A.; Heghinian, M. D.; Sahil, T.; Adams, D. J.; Marí, F.; Craik, D. J. Transforming conotoxins into cyclotides: Backbone cyclization of P-superfamily conotoxins. *Biopolymers* **2015**, *104*, 682–692.