Optimization of a β-sheet-cap for long loop closure

Jordan M. Anderson, Alexander A. Shcherbakov, Brandon L. Kier, Jackson Kellock, Irene Shu, Aimee L. Byrne and Niels H. Andersen*

Department of Chemistry, University of Washington, Seattle WA 98195

Supporting Information.

All relevant ^{1}H NMR data used to produce this manuscript is stored in a publicly accessible Chemical Shift Database (CSDb). In order to access, search for the CSDb ascension number or name at the website: http://andersenlab.chem.washington.edu/CSDb/. In addition to acting as a depository for this data, the website contains the algorithm[1–4] used to calculate CSD values and contains a graphical program for comparing histograms of different peptides.

Table 1. Peptide List

Entry #	System	Loop Length	**Sequence** cap – β-strand – loop – β-strand – cap’	Ref.	CSDb Ascension Name
1		10	**Ac-W**VTI - G_3GKKGG_3 - KKI**WTG**-NH_2	603	GLHP
2		10	**Pr-W**VCK - G_3GKKGG_3 –VCK**WTG**PK-NH_2	703	Superloop
3	B2	2	**Ac-W**ITVTI - GG - KKIRV**WTG**-NH_2	780	cap@6 GG
4	B2	4	**Ac-W**ITVTI – GGGG - KKIRV**WTG-NH_2**	784	GGGG
19	B2	6	**Ac-W**ITVTI-GGGGGG-KKIRV**WTG-NH_2**	870	Cap @ 6c G6
5	B2	10	**Ac-W**ITVTI - G_3GKKGG_3 - KKIRV**WTG**-NH_2	739	cap@6 superloop
20	B2	6	**Ac-W**ITVTI-(G**IpGK**G)-KKIRV**WTG-NH_2**	806	GIpGKG
22	B2	10	**Ac-W**ITVTI-(G_3**IpGK**G_3)-KKIRV**WTG-NH_2**	815	3G-IpGK
21	B2	8	**Ac-W**ITVTI-(G**NPATGK**G)-KKIRV**WTG-NH_2**	808	GNPATKG
6	B4	10	**Ac-W**ITVRI**W**-G_3GKKGG_3-**W**KTIRV**WTG**-NH_2	833	W4G4K2G4
7	B4	2/4	**Ac-W**ITVRI**W** - SNGK - **W**KTIRV**WTG**-NH_2	824	W4SNGK
8	C2 R1H	10	**HW**ITVTI - G_3GKKGG_3 – KKIRV**WE**	803	Superloop HWWE
9	C2	2	**RW**ITVTI - GG - KKIRV**WE**	984	R-GG-E
10	C2	4	**RW**ITVTI - GGGG - KKIRV**WE**	901	RGGGGE

23	C2	6	**RW**ITVTI – GGGGGG - KKIRV**WE**	1065	R-GGGGGG-E
11	C2	10	**RW**ITVTI - $(G_3GKK)GG_3$ - KKIRV**WE**	922	R-Superloop-E
12	C2	16	**RW**ITVTI - $(G_3GKK)_2GG_3$ - KKIRV**WE**	951	R-(GGGGKK)3-E
13	C2	22	**RW**ITVTI - $(G_3GKK)_3GG_3$ - KKIRV**WE**	967	R-(GGGGKK)4-E
24	C2	6	**RW**ITVTI - G**IpGK**G - KKIRV**WE**	1095	R-GIpGK-E
25	C2	6	**RW**ITVTI - G**IPGK**G - KKIRV**WE**	1109	R-GIPGKG-E
26	C2	8	**RW**ITVTI - G**NPATGK**G - KKIRV**WE**	1112	R-GNPATGKG-E
27	C2	10	**RW**ITVTI - G_3**IpGK**G_3 - KKIRV**WE**	1094	R-G3IpGKG3-E
28	C2	10	**RW**ITVTI - G_3**IPGK**G_3 - KKIRV**WE**	1122	R-G3IPGKG3-E
29	C2	16	**RW**ITVTI - G_6**IpGK**G_6 - KKIRV**WE**	1106	R-G6IpGKG6-E
30	C2	16	**RW**ITVTI - G_6**IPGK**G_6 - KKIRV**WE**	1127	R-G6IPGKG6-E
31	C2	16	**RW**ITVTI – G_5**NPATGK**G_5 - KKIRV**WE**	1178	R-GNPATGKG5-E
14	C4	2/4	**RW**ITVRI**W** - IGGK - **W**KTIRV**WE**	1364	RWW-IGGK
15	C4	10	**RW**ITVRI**W** - G_3GKKGG_3 - **W**KTIRV**WE**	1055	RWW-10loop-WWE
16	C4	16	**RW**ITVRI**W**- $(G_3GKK)_2GG_3$ -**W**KTIRV**WE**	1156	RWW-16Loop-WWE
16Y		16	**RW**ITVRI**Y** - $(G_3GKK)_2GG_3$ -**W**KTIRV**WE**	1409	RWY-16loop-WWE
17	C4	22	**RW**ITVRI**W**- $(G_3GKK)_3GG_3$ -**W**KTIRV**WE**	NMR data not available.	
18	C4	28	**RW**ITVRI**W**- $(G_3GKK)_4GG_3$ -**W**KTIRV**WE**	NMR data not available.	

Fraction Folded Calculation (χ_F)

R W I T V T I – loop – K K I R V W E
S-7 S-5 S-3 S-1 S+1 S+3 S+5 S+7

Fraction folded populations were calculated using an average of CSDs from backbone Hαs S-6, S-4, S+4 and H_Ns S-5, S-3, S+3, S+5 with additional probes from the sidechain of the Tryptophan in the edge position, Hβ3, Hζ3, and Hε3 for each Trp/Trp aryl cluster. Additionally, systems B2 and B4 also include the diagnostic C-terminal Gly H_N.

Figure S1. Backbone H_N CSDs of the peptides used as fully folded baselines, at 280-320K. Little to no melting is observed upon heating.

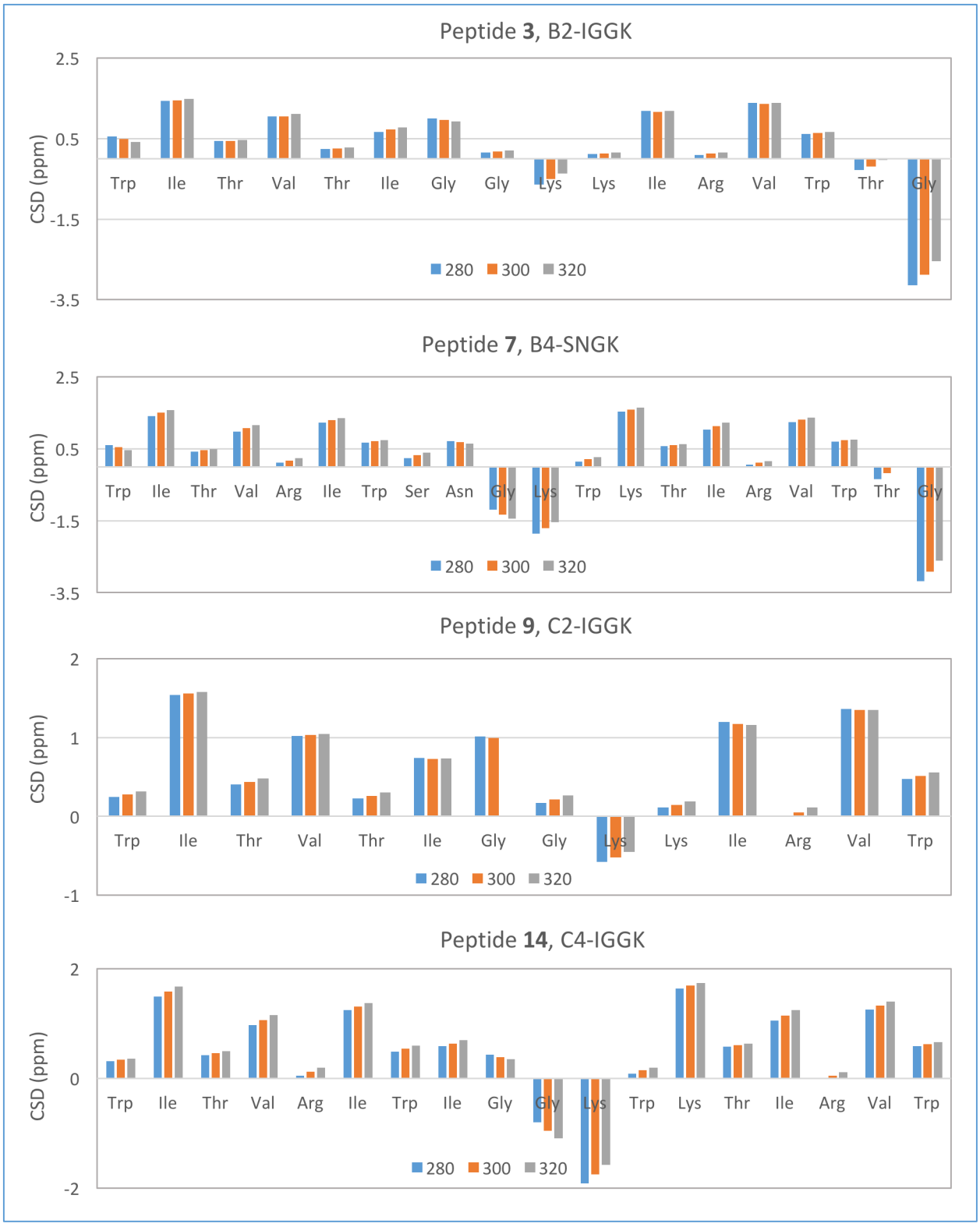


Figure S2. RWITVRIW^{Flank} – Loop – WKTIRVW^{Cap}E Comparison of the position specific fraction folded of the termini and residues near the loop.

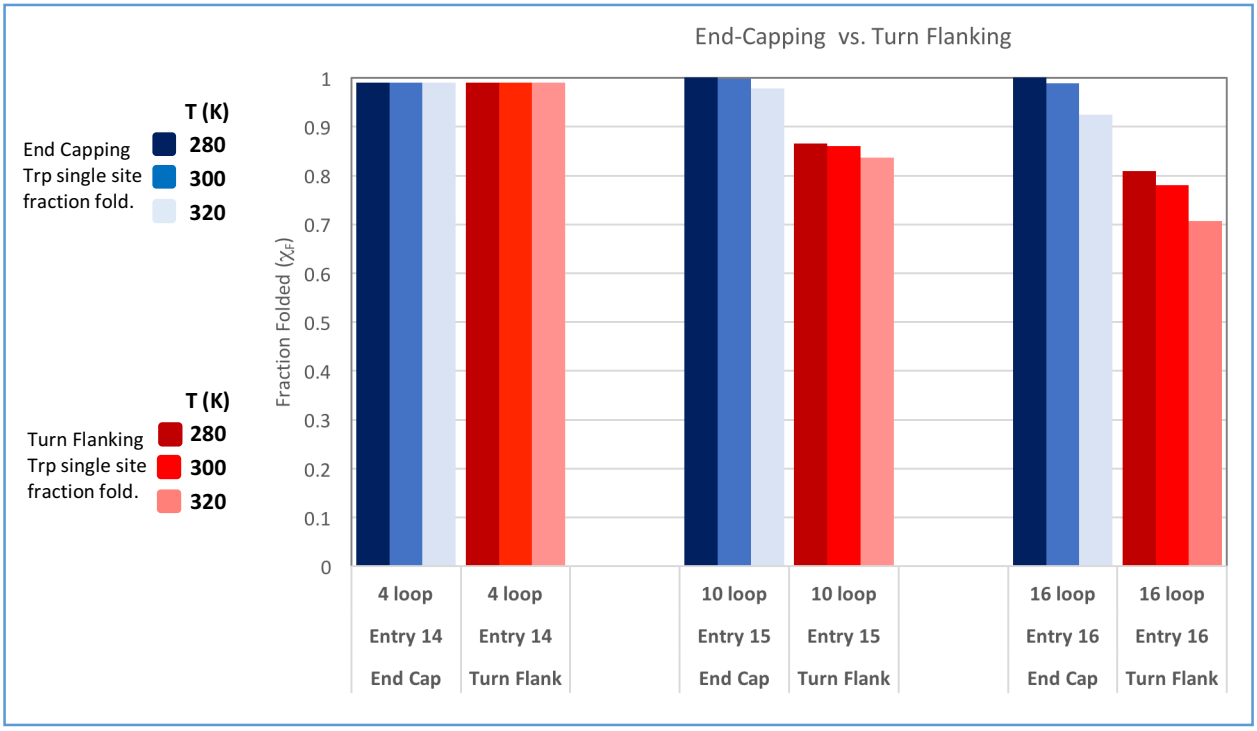


Figure S3. CD spectra was taken from 6.25-100 μM to verify the exciton couplet observed is from a monomeric species. (A) Raw CD spectra. (B) Signal at 228 nm vs. Peptide concentration.

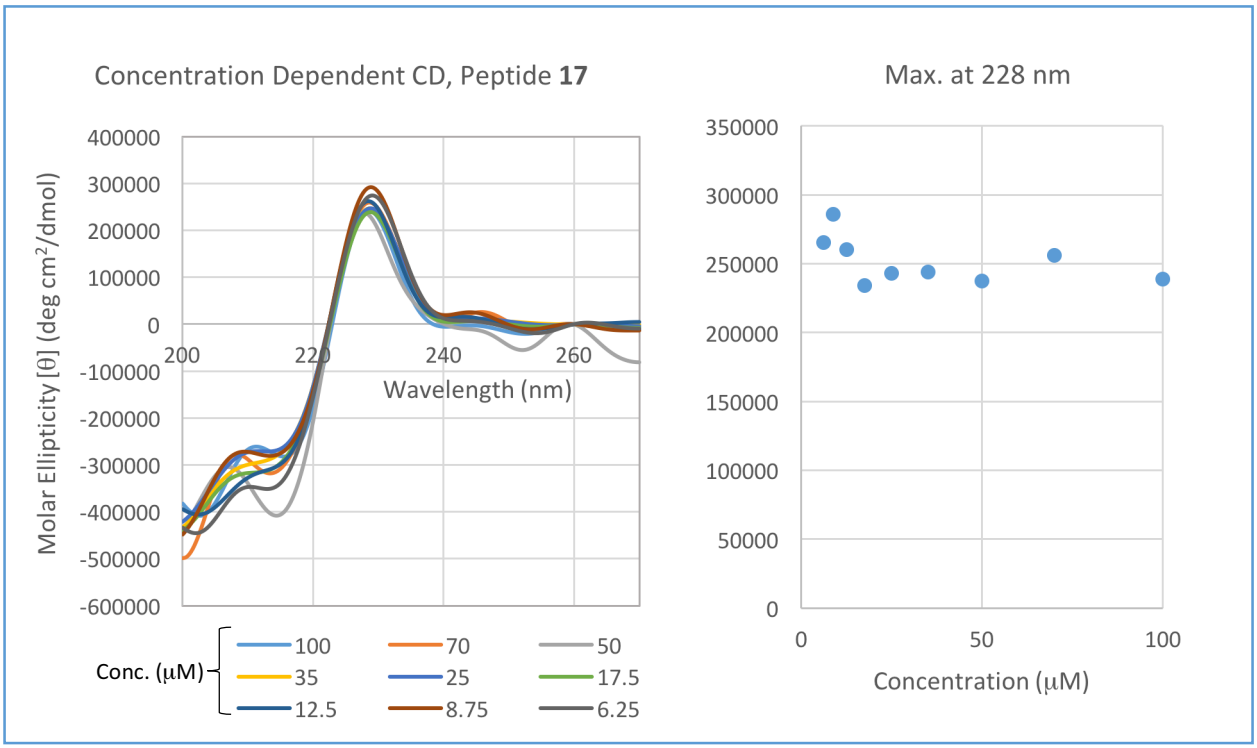


Figure S4. Concentration dependent NMR spectra of Peptide **17** (system C4-22 loop): NMR spectra of the aromatic and amide H_N region, taken at 30 (red) and 700 μM (blue). Peak position and broadening appear similar. The broadening seen, particularly at highly shifted resonances, is due to slow exchange between the folded and unfolded states.

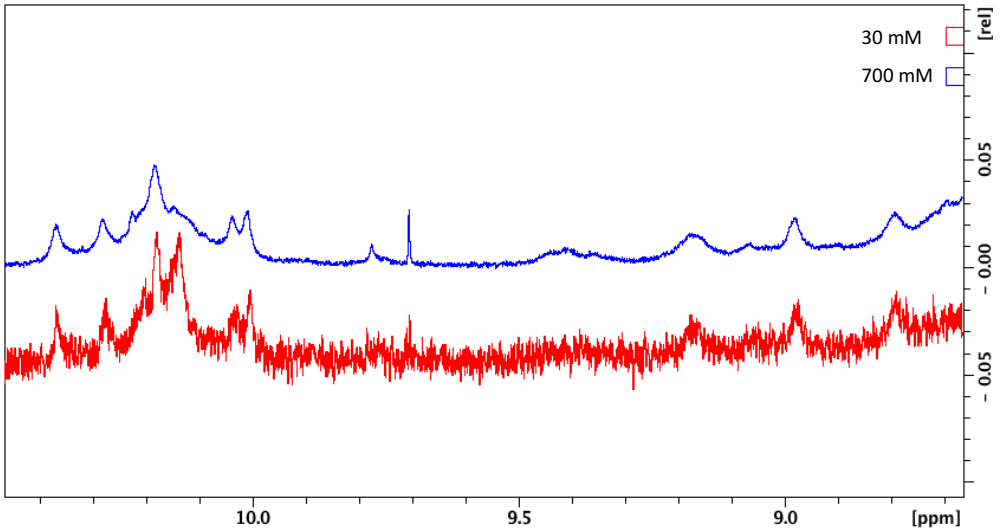
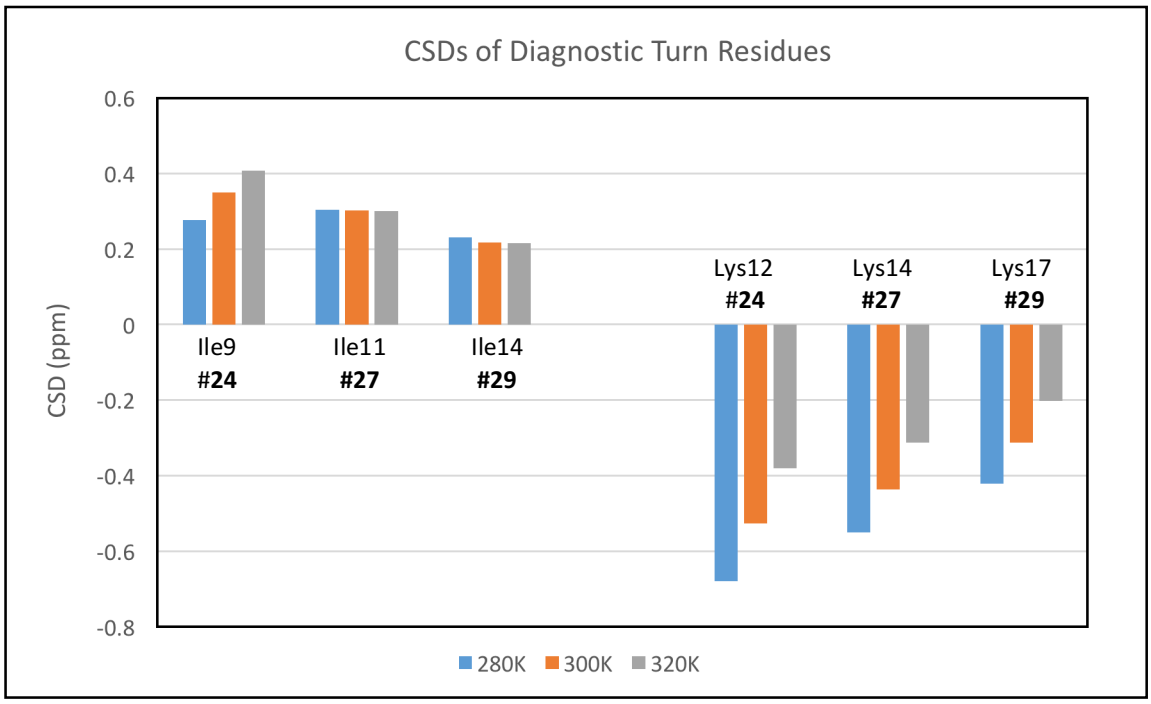


Figure S5. CSDs of the diagnostic probes within the added turn-region (H_N of Ile and Lys in the turn, underlined in each sequence) of different length loops.

Peptide **24:** RWITVTI-GIpGKG-KKIRVWE
Peptide **27:** RWITVTI-GGGIpGKGGG-KKIRVWE
Peptide **29:** RWITVTI-GGGGGGIpGKGGGGGG-KKIRVWE



References

(1) Fesinmeyer, R. M., Hudson, F. M., and Andersen, N. H. (2004) Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Hairpin. *J. Am. Chem. Soc. 126*, 7238–7243.

(2) Andersen, N. H., Cao, B., and Chen, C. (1992) Peptide/Protein Structure Analysis Using the Chemical Shift Index Method: Upfield α-CH Values Reveal Dynamic Helices and α_L Sites. *Biochem. Biophys. Res. Commun. 184*, 1008–1014.

(3) Andersen, N. H., and Tong, H. (1997) Empirical parameterization of a model for predicting peptide helix/coil equilibrium populations. *Protein Sci. 6*, 1920–1936.

(4) Eidenschink, L., Kier, B. L., Huggins, K. N. L., and Andersen, N. H. (2009) Very short peptides with stable folds: Building on the interrelationship of Trp/Trp, Trp/cation, and Trp/backbone-amide interaction geometries. *Proteins Struct. Funct. Bioinforma. 75*, 308–322.