Electronic Supplementary Information

Gas phase folding of a $(Ala)_4$ neutral peptide chain: spectroscopic evidence for the formation of a β-hairpin H-bonding pattern

by E. Gloaguen, R. Pollet, F. Piuzzi, B. Tardivel, M. Mons*

CEA, IRAMIS, Service Photons Atomes et Molécules, F-91191 Gif-sur-Yvette, France.
CNRS, LFP, Biomolecular Structures Group, F-91191 Gif-sur-Yvette, France.

Appendix S1

The conformational exploration of the Ac-$(Ala)_4$-O-Bzl system has been carried out with AMBER99 [1] and Charmm22 [2] force fields using the HyperChem 7.52 Package.[3] After a trial on the number of H-bonds (3 or 4), 55 conformations over 2000 optimized geometries at the force field level were selected for an optimization at the DFT-D level (B97-D/TZVPP) using the TurboMole 5.10 package.[4].

In order to check the robustness of the conformations capable to account for experimental data (3 strong H-bonds), the same force field structures were also optimized at the MP2/SVP level, which is also expected to account for dispersion, although less accurately owing to the modest basis set imposed by our computer facility. It turned out that additional minima were obtained at this level. In order to get an extended and consistent view of the potential energy landscape, these additional minima were re-optimized at the DFT-D level of theory. For selected forms with stable 3 H-bonds conformations, possible positions of the floppy tail were investigated at a force field level, after freezing the backbone geometry.

The resulting DFT-D optimizations has eventually given rise to 25 configurations of comparable energetics (within less than 3.3 kcal mol^{-1}).

References

1. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., *J. Comput. Chem.* **2005,** 26, 1668-1688.
2. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., *J. Phys. Chem. B* **1998,** 102, 3586-3616.
3. *HyperChem* 7.52; Hypercube: 2005.
4. Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C., *Chem. Phys. Lett.* **1989,** 162, 165-169.

Table S1

Table S1: Conformations of Ac-$(Ala)_4$-O-Bzl, as obtained after potential surface exploration and subsequent energy optimization at the RI-B97-D/TZVPP level of theory, sorted according to their H-bond network. 0 K enthalpies and 300 K free energy are given in kcal mol^{-1}. Free energies have been estimated from the chemical potentials using the Turbomole package,[4] assuming a rigid rotator and negligible vibrational anharmonicity effects for the calculation of partition functions. The Ramachandran dihedral angles of each conformation are given in degrees. Scaled (0.9744) NH strech harmonic frequencies are given in cm^{-1}.

Conformation		ΔH	ΔG	Ramachandran dihedral angles								Vibrations			
		0 K	300 K	Ala (1)		Ala (2)		Ala (3)		Ala (4)		NH stretch frequencies			
type		kcal/mol	kcal/mol	φ	ψ	φ	ψ	φ	ψ	φ	ψ	1	2	3	4
Experiment	***A***											***3269***	***3332***	***3383***	***not obs'd***
	B											***3289***	***3353***	***3383***	***not obs'd***
14-7L-x-10(II')	**1β-a**	0.23	**0.00**	-81	75	56	-121	-84	-10	-84	167	**3294**	**3332**	**3374**	**3475**
	1β-b	1.25	1.73	-82	72	56	-122	-73	-9	-96	68	**3278**	**3332**	**3375**	**3472**
	1β-c	2.53	1.30	-81	76	55	-122	-81	-12	-81	151	**3289**	**3330**	**3363**	**3477**
14-7L-[7D-10(II')]	**1m1-b**	**0.00**	0.62	-83	67	67	-88	-96	-9	-112	87	3233	3326	3401	3440
	1m2-b	1.97	3.75	-81	71	68	-73	-102	-16	-130	75	3191	3330	3365	3370
14-7L-7D-x	**1γ -a**	1.72	1.88	-83	65	68	-66	-136	23	-147	173	3276	3304	3401	3465
14-7D-[7L-10(II)]	**1'm-a**	1.98	4.42	72	-61	-66	107	51	37	-160	52	3209	3409	3432	3444
14-[7L-10(II)]-10(I')	**2m-a**	0.15	1.18	-77	97	52	42	77	-12	-77	132	3297	3397	3409	3419
	2m-b	0.92	1.39	-78	90	52	43	76	-11	-90	178	3297	3395	3412	3443
14-7L-x-10(I')	**2β -a**	2.67	2.87	-72	108	52	47	84	-26	-154	-81	3371	3377	3440	3448
14-7D-x-10(I)	**2'γ-a**	2.07	2.40	73	-77	-60	-31	-103	28	57	-136	3261	3292	3408	3453
14-X-10(II')-10(I)	**2'β -a**	3.01	4.77	61	-107	-54	-30	-90	-16	-177	114	3331	3345	3374	3457
14-[7D-10(II')]- 10(I)	**2'm-a**	0.95	1.80	70	-87	-60	-31	-102	26	67	-120	3303	3386	3407	3418
	2'm2-b	1.74	3.23	65	-93	-61	-27	-97	31	84	151	3371	3420	3424	3429
	2'm3-c	2.05	2.20	70	-83	-83	-59	-100	30	56	128	3269	3355	3410	3421
x-7L-7L-7L	**3A**	2.57	2.25	-82	78	-82	62	-81	76	52	35	3281	3326	3346	3469
x-7L-7D-7L	**3B**	3.30	3.70	-83	63	69	-73	-85	63	168	7	3276	3280	3369	3464
π-7D-7L-7D	**4A**	0.06	1.95	72	-50	-80	73	70	-66	-74	131	3185	3248	3254	3468
14-7D-7L-7D	**4B**	1.77	2.45	71	-56	-76	86	72	-64	-50	145	3234	3286	3331	3443
π-7L-7D-7L	**4'A-a**	1.21	2.92	-82	77	71	-58	-82	53	49	38	3250	3279	3289	3429
π -7L-7D-7L	**4'A2-b**	1.51	3.30	-83	64	69	-61	-81	76	60	-137	3211	3253	3298	3429
14-7L-7D-7L	**4'B**	1.65	2.80	-82	64	37	-73	-84	61	54	-154	3271	3282	3326	3405
π -11-10(I)-7D	**5A**	0.07	0.83	-64	-27	-124	39	70	-65	-71	165	3247	3326	3439	3444
5-11-7D-7L	**5B**	1.85	1.51	-99	115	-84	71	67	-69	-87	162	3301	3343	3446	3465
x-x-10(I) -10(I)	**6**	2.90	1.27	-66	-24	-71	-7	-76	-15	-101	-17	3402	3413	3424	3429

Fig. S1 : DFT-D optimized structures (B97-D/TZVPP level of theory) of the most stable conformations of Ac-$(Ala)_4$-O-Bzl molecule found in the exploration described in Appendix S1. H-bond distances are given in pm. The ZPE-corrected relative energies (brackets) are given in kcal mol^{-1}. The notation indicates the H-bonding status of each of the 4 NH's along the chain; x stands for a free NH; elongated bonds (mixed structures) are indicated between brackets [ ]. For the C_{10} H bonds, the type of β-turn (I, II, I' or II') is also indicated. Labels refer to the 6 families described in the paper: β and γ refer to the presence of a β- and γ-turn respectively, m to a mixed behavior; capital letters indicate different backbones and lower case letters indicate different arrangements of the Ala-O-Bzl tail.

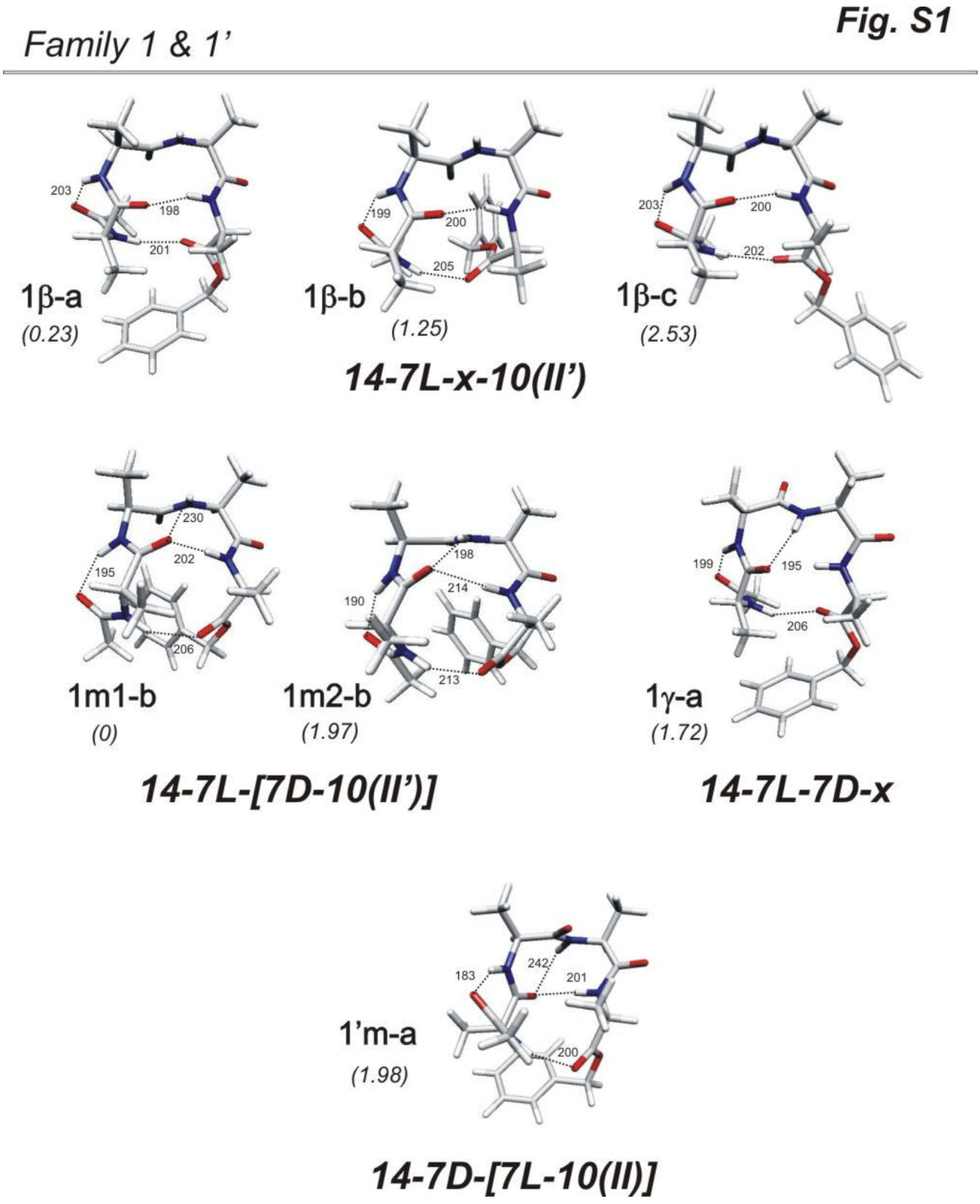
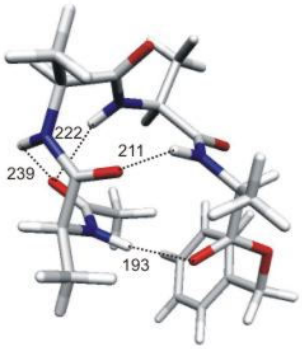
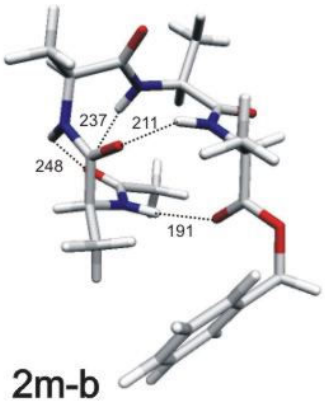


Fig. S1 (cont'd)

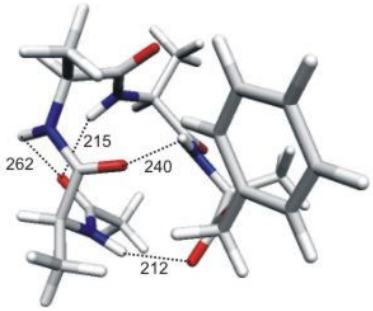
Family 2 & 2'



2m-a
(0.15)

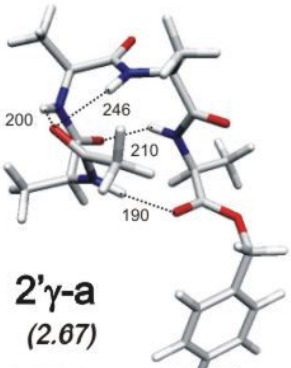


2m-b
(0.92)

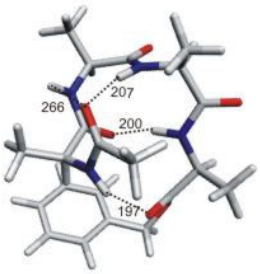


2β-a
(2.67)

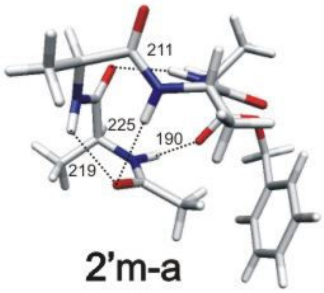
14-[7L-10(II)]-10(I') ***14-x-10(II)-10(I')***



2'γ-a
(2.67)

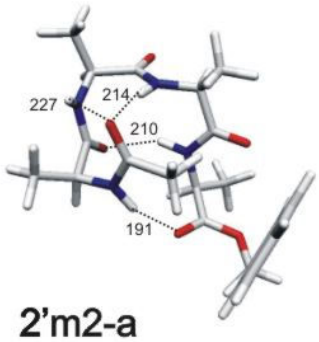


2'β-a
(3.01)

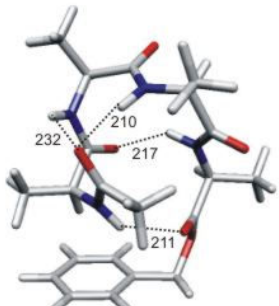


2'm-a
(0.95)

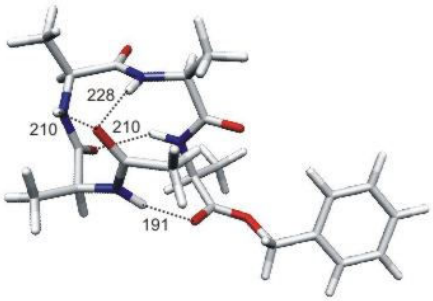
14-7D-x-10(I) ***14-x-10(II')-10(I)*** ***14-[7D-10(II')]-10(I)***



2'm2-a
(1.02)



2'm3-a
(1.74)

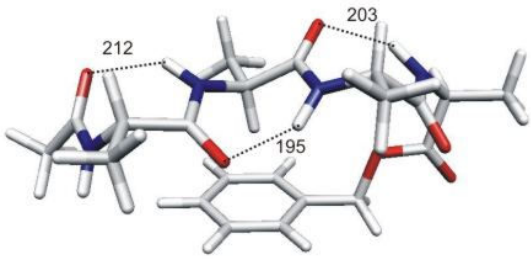


2'm4-a
(2.05)

14-[7D-10(II')]-10(I)

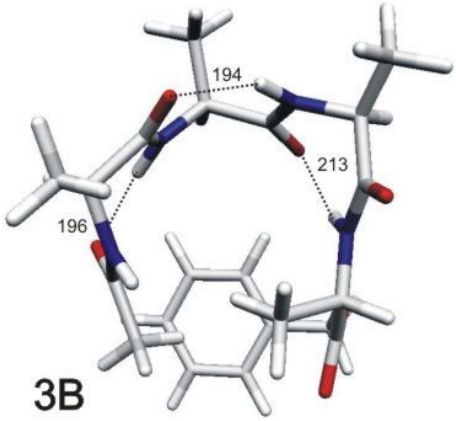
Family 3

Fig. S1 (cont'd)



3A *x-7L-7L-7L*

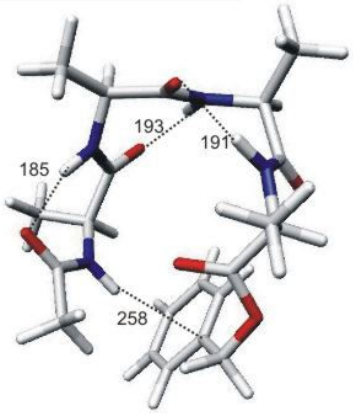
(2.57)



3B *x-7L-7D-7L*

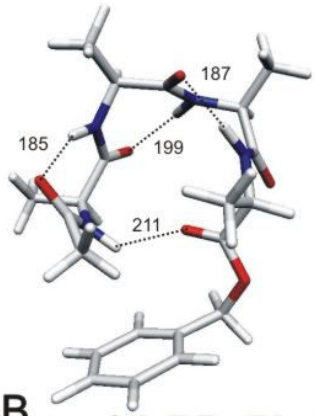
(3.30)

Family 4 & 4'



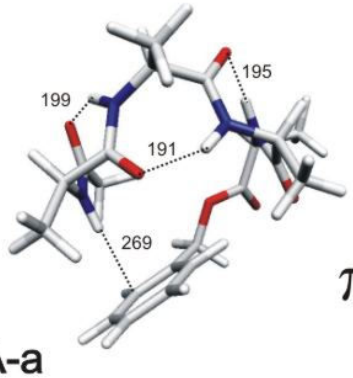
4A *π-7D-7L-7D*

(0.06)



4B *14-7D-7L-7D*

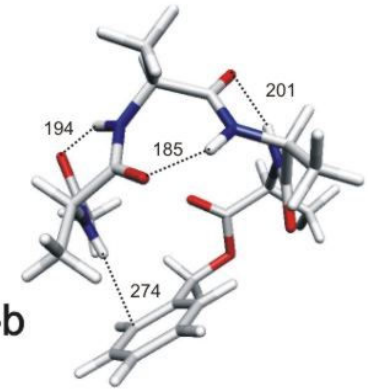
(1.77)



π-7L-7D-7L

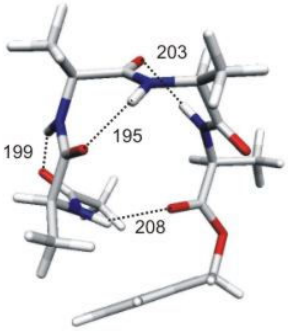
4'A-a

(1.21)



4'A2-b

(1.51)

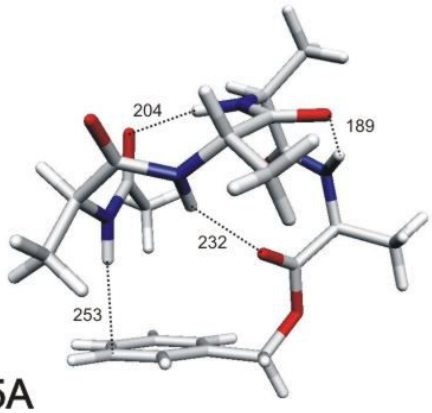


4'B *14-7L-7D-7L*

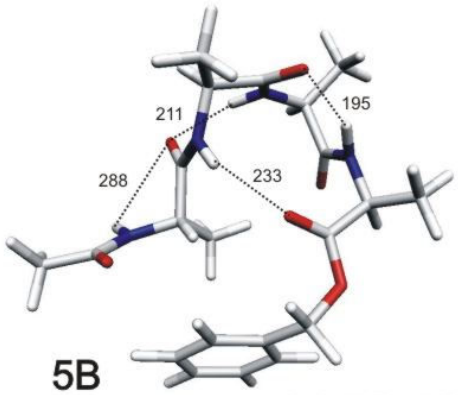
(1.65)

Family 5

Fig. S1 (end)

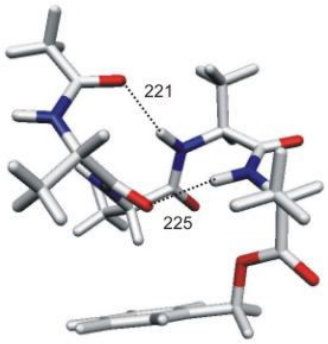


5A
(0.07) **π-11-10(l)-7D**



5B
(1.85) **5-11-7D-7L**

Family 6



6
(2.90) ***x-x-10(l)-10(l)***