Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Unifying the microscopic picture of His-containing turns: from gas phase model peptides to crystallized proteins

Woon Yong Sohn†, Sana Habka, Eric Gloaguen and Michel Mons*

LIDYL, CEA, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette Cedex, France

Supplementary information

Table of content

Figure S10 : Structure of the most stable monohydrate of the ε-6^{δ} His His conformation of the Ac-His-NH_2 model in a γ-turn backbone conformation.

Figure S11: Structure of the two most stable monohydrates of the δ-$^{\delta}7/\pi^{H}$ His conformation of the Ac-His-NH_2 model in a γ-turn backbone conformation.

Table S1 : Z-His-NH_2 molecule : Experimental NH stretch frequencies (cm^{-1}) of the three conformers (A-C) detected in the supersonic expansion, compared to the predicted frequencies of conformations of various backbone types for each of the His side chain tautomers (ε vs. δ), as given in Fig. 1. Also given are, the energetics of these conformations (kJ/mol), two indicators of the error between predicted spectra and observed conformers' frequencies (δ^{ave} , the average RMS error the unsigned maximum error). The best fits according to frequency (small δ^{ave} and δ^{max} values) and energetic criteria (small ΔG values) are indicated in bold. Predicted frequencies are obtained at the B97-D3/TZVPP level of theory, by applying a mode-dependent two-parameter scaling factor[b] to the harmonic NH stretch frequencies.[1] For all these molecules, only the most stable form of the benzylcarboxy (Z-) cap, i.e., its extended conformation (N(i)C(i+1)$OC_{carb.}$ ~180°), has been considered.

	experimental										
	A				**3172 ; 3410 ; 3505 ; 3514**						
	B				**3333 ; 3360 ; 3515 ; 3520**						
	C				**3349 ; 3398 ; 3420 ; 3516**						
Conformation short and full names			ΔH	ΔG 300 K	NH_2^{anti}	NH_2^{sym}	NH_{His}	$NH_{sidechain}$	A $\delta^{ave}/\delta^{max}$	B $\delta^{ave}/\delta^{max}$	C $\delta^{ave}/\delta^{max}$
δ1	**π^{H}-7 ($^{\delta}$7)**	**C**	**0**	**0**	**3515**	**3338**	**3416**	**3404**	65/166	38/98	**6/11**
ε1	**5-7^{δ} ($^{\varepsilon}$f)**	**A**	**6**	**1**	**3506**	**3199**	**3405**	**3504**	**10/27**	51/134	63/150
δ2	5-π^{H} ($^{\delta}$8)		9	4	3532	3403	3439	3410	79/231	52/75	25/54
ε2	**6^{δ}-7($^{\varepsilon}$f)**	**B**	**8**	**5**	**3515**	**3349**	**3343**	**3501**	59/171	**10/13**	33/81

Table S2 : Ac-Phe-His-NH_2 molecule : Experimental NH stretch frequencies (cm^{-1}) of the four conformers (A-D) detected in the supersonic expansion, compared to the predicted frequencies of the conformations of various backbone types for each of the His side chain tautomers (ε vs. δ). Same details as in Table S1. Conformers A and D are well accounted for by the lowest ε-type conformations. For conformers B and C, which both belong to the δ-type (owing to the absence of spectral signature of a free NH_{His} group), three low energy conformations, δ1, δ2 and δ4, can match frequencies. δ1 matches both B and C, but is assigned to B, by analogy in terms of UV spectroscopy with the assignment with a very similar system (the B conformer of Ac-Phe-Phe-NH_2).[2] δ2 is assigned to C on an energetic basis. δ4 which could account for both B and C is higher in energy; it differs from δ2 only by the orientation of the Phe side chain (*gauche-* vs. *gauche+*), and is expected to relax easily to the lowest d1 form, due to a modest barrier to rotameric isomerization.

	experimental										
	A				**3357 ; 3385 ; 3440 ; 3511 ; 3523**						
	D				**3265 ; 3329 ; 3442 ; 3511[a]**						
Conformation short and full names			ΔH	ΔG 300 K	NH_2^{anti}	NH_2^{sym}	NH_{Phe}	NH_{His}	$NH_{sidechain}$	A $\delta^{ave}/\delta^{max}$	D[a] $\delta^{ave}/\delta^{max}$
ε1	**π^F-6$^\delta$-10 ($^\varepsilon$f)**	**A**	**8**	**3**	**3527**	**3393**	**3444**	**3356**	**3502**	**5/9**	36/91
ε2	**5-6$^\delta\pi^F$-7 ($^\varepsilon$f)**	**D**	**10**	**7**	**3515**	**3334**	**3443**	**3244**	**3504**	38/113	**8/21**
ε3	9$^\delta\pi^F$-6$^\delta$-7 ($^\varepsilon$f)		10	8	3522	3368	3298	3452	3510	22/59	19/39
ε4	5-6$^\delta\pi^F$-π^F($^\varepsilon$f)		13	8	3546	3433	3447	3322	3501	28/52	42/108
ε5	5-5π^F-7$^\delta$ ($^\varepsilon$f)		17	9	3503	3199	3442	3366	3504	41/158	24/66
ε6	5-6$^\delta\pi^F$-7 ($^\varepsilon$f)		15	9	3503	3337	3430	3224	3504	44/133	15/41
ε7	5-6$^\delta\pi^F$-7 ($^\varepsilon$f)		18	10	3510	3338	3443	3404	3500	13/19	32/75
ε8	π^F-6$^\delta$-7($^\varepsilon$f)		15	11	3513	3326	3448	3325	3501	24/59	16/60
ε9	5-6$^\delta\pi^F$-7 ($^\varepsilon$f)		14	13	3529	3404	3438	3387	3506	12/30	45/122
ε10	9$^\delta$-f-f ($^\varepsilon$f)		13	13	3551	3436	3167	3423	3508	53/190	42/118
ε11	5-6$^\delta$-7 ($^\varepsilon$f)		21	16	3499	3351	3441	3352	3501	15/33	49/98
	experimental										
	B				**3326 ; 3377 ; 3384 ; 3445 ; 3508**						
	C				**3351 ; 3372 ; 3396 ; 3450 ; 3511**						
	Conformation		ΔH	ΔG 300 K	NH_2^{anti}	NH_2^{sym}	NH_{Phe}	NH_{His}	$NH_{sidechain}$	B $\delta^{ave}/\delta^{max}$	C $\delta^{ave}/\delta^{max}$
δ1	**π^F-π^H-7 ($^\delta$7)**	**B**	**0**	**0**	**3514**	**3350**	**3443**	**3389**	**3394**	**11/24**	**8/17**
δ2	**5-$\pi^{F/H}$-7 ($^\delta$7)**	**C**	**3**	**1**	**3517**	**3354**	**3449**	**3375**	**3414**	**15/30**	**9/18**
δ3	π^F-π^H-10 ($^\delta$f)		0	3	3523	3388	3437	3418	3446	34/62	30/50
δ4	π^F-π^H-7 ($^\delta$7) [c]		9	7	3507	3318	3456	3383	3387	**6/11**	**10/20**

a) For assignment of conformer D, the band at 3511 cm^{-1} has been considered as due to a spectral coincidence of two bands.

b) $freq^{scaled} = a * freq^{harm} + b$ with (a; b) values of (1.03967; -229.94 cm^{-1}), (0.90447; 253.58 cm^{-1}) and (0.76584; 751.22 cm^{-1}) for, respectively, the NH stretch modes of amide and imidazole; the symmetric NH_2 stretch mode and the antisymmetric NH_2 stretch mode.

c) Differs from δ_1 by the Phe side chain orientation

Table S3 : Ac-His-Phe-NH_2 molecule : Experimental NH stretch frequencies (cm^{-1}) of the major conformer A detected in the supersonic expansion, compared to the predicted frequencies of the first conformations of various backbone types for each of the His side chain tautomers (ε vs. δ). Same details as Table S1.

	experimental									
	A				***3355 ; 3375 ; 3442 ; 3510 ; 3522***					
	conformation		ΔH	ΔG (300 K)	NH_2^{anti}	NH_2^{sym}	NH_{Phe}	NH_{His}	$NH_{sidechain}$	A $\delta^{ave}/\delta^{max}$
ε1	**$6^{\delta-}\pi^{F/am1}$-10($^{\varepsilon}$f)**	**A**	**0**	**0**	**3523**	**3380**	**3456**	**3348**	**3499**	8/14
ε2	6^{δ}-π^{am1}-10($^{\varepsilon}\pi^{F}$)		2	5	3518	3373	3462	3338	3444	15/48
ε3	5-$\pi^{H/F}$-10^{δ}($^{\varepsilon}$f)		5	5	3493	3201	3447	3426	3504	49/154
ε4	5-7^{δ}-10^{δ}($^{\varepsilon}$f)		20	13	3501	3275	3431	3417	3499	33/80
δ1	π^{HF}-7-π^{am2}($^{\delta}$7)		6	6	3555	3440	3231	3394	3414	55/124
δ2	π^{H}-$\pi^{F/am1}$-10($^{\delta}$7)		7	10	3527	3398	3461	3421	3449	30/49

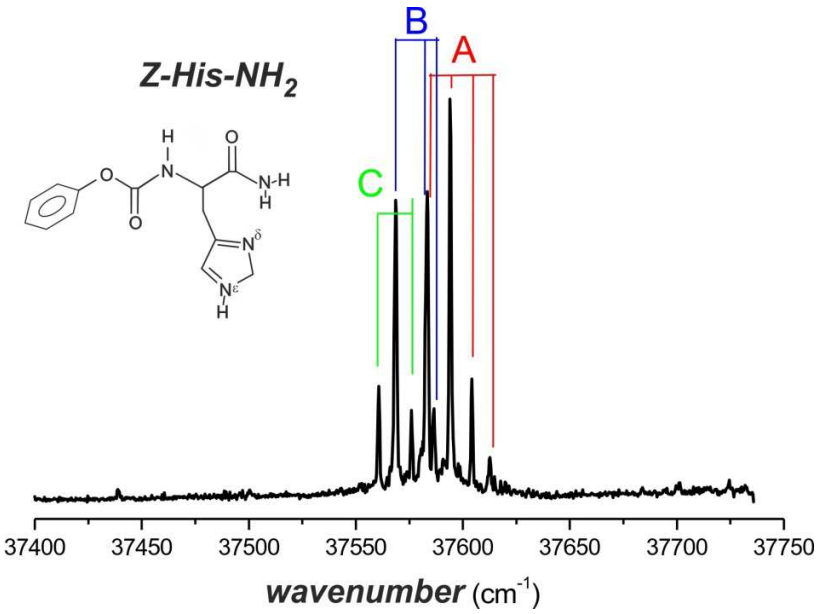

Figure S1 : Mass-resolved near UV spectrum of jet-cooled Z-His-NH_2, in the origin region of the first electronic transition of the phenyl ring, as obtained by resonant two-photon ionization spectroscopy. The IR/UV double resonance spectra carried out by probing each of the bands observed has allowed us to determine the existence of 3 conformers and the corresponding three sets of UV transitions.

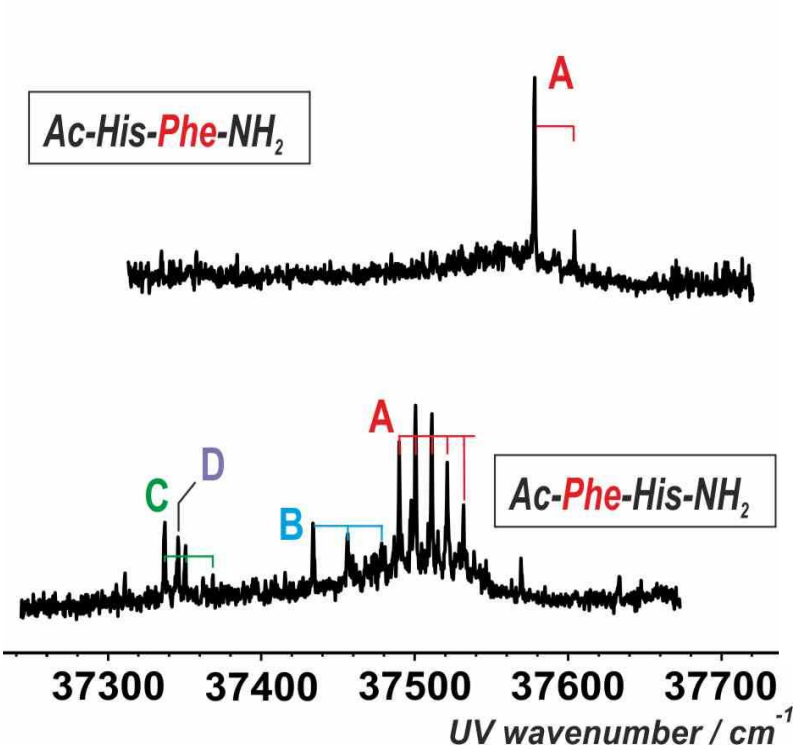

Figure S2 : Mass-resolved near UV spectrum of the jet-cooled Ac-Phe-His-NH_2 and Ac-His-Phe-NH_2 molecules, in the origin region of the first electronic transition of the phenyl ring, as obtained by resonant two-photon ionization spectroscopy. The IR/UV double resonance spectra carried out by probing each of the bands observed has allowed us to determine the existence of one and four conformers in the top and bottom spectra respectively.

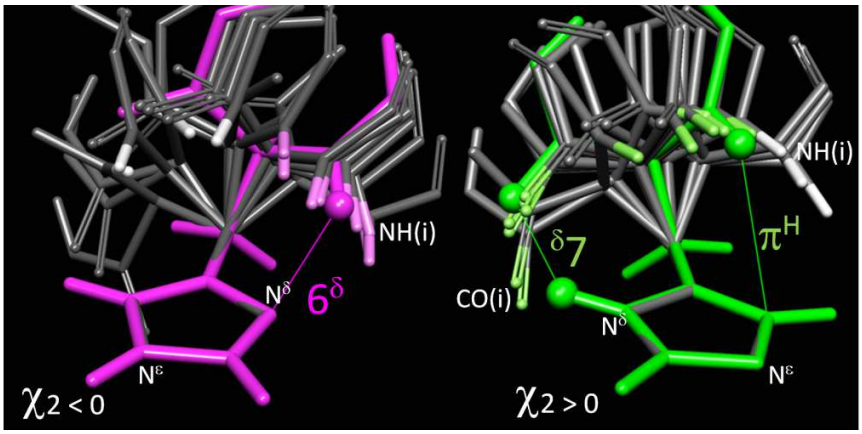

Figure S3: Backbone distribution around a His side chain in the i+1 position of a γ-turn (g+ rotamer), obtained from selected His residues in the PDB, sorted according to their χ_2 value. a) left: χ_2– subset, showing the backbone NH(i) moieties (in light magenta) distributed in the vicinity of the nearest δ atom of the ring, compared to the gas phase ε-6^{δ} conformations of the Ac-Phe-His-NH_2 and Ac-Ala-His-NH_2 models; b) right: χ_2+ subset, showing in light green the backbone NH(i) moieties engaged in a π^H H-bond, compared to the gas phase δ-($^{\delta}7/\pi^H$) conformations of the Ac-His-NH_2 model highlighted in green, with H(i) and O(i) atoms figured as green balls. NH(i) moieties which do not fit gas phase structures are left in white.

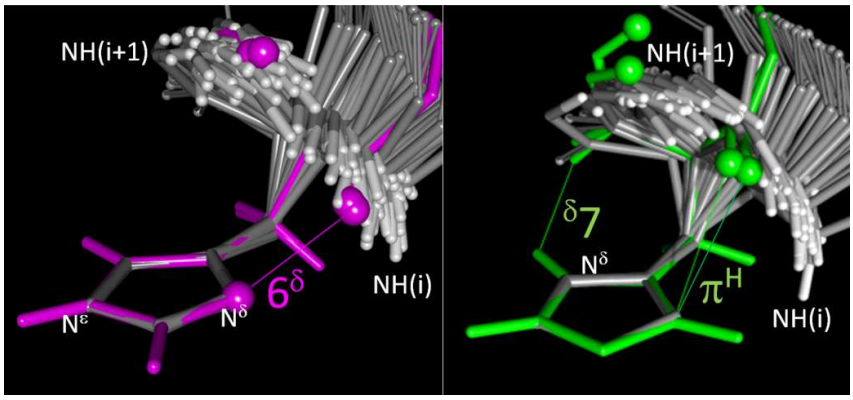

Figure S4: Backbone distribution around His residues in the i+1 position of a β-turn, obtained from selected His residues in the PDB, sorted according to their χ_2 value. a) left: χ_2- subset, showing in white the backbone NH(i) moieties in the vicinity of the nearest δ atom of the ring, compared to the gas phase ε-6^{δ} conformations of the Ac-Phe-His-NH_2 and Ac-Ala-His-NH_2 models. In this case the H(i) atoms of the gas phase models (lower magenta balls) are located on the edge of the NH(i) protein distribution. b) right: χ_2+ subset, showing the distribution of backbone NH(i) moieties (in white), which does not fit to the gas phase δ-$\pi^H/^{\delta}7$ conformation of the Ac-Phe-His-NH_2 and Ac-Ala-His-NH_2 models (δ tautomer) stabilized by $^{\delta}7$ and π^H H-bonds (structures highlighted in green).

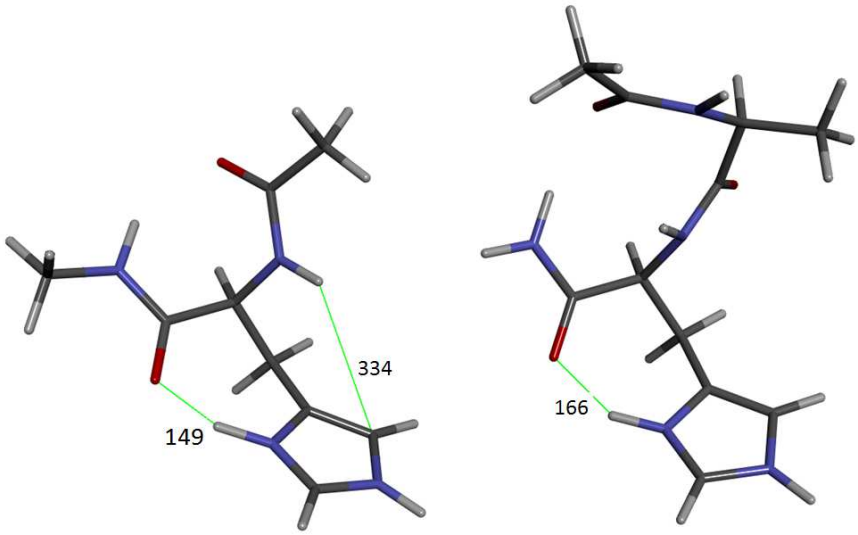

Figure S5: Optimized structures (RI-B97-D3/TZVPP level of theory) of γ- and β-turns in the protonated Ac-His(H^+)-NH-Me and Ac-Ala-His(H^+)-NH_2 model molecules, showing a shortened $^{\delta}7$ H-bond. All distances are in pm, in this figure and the following.

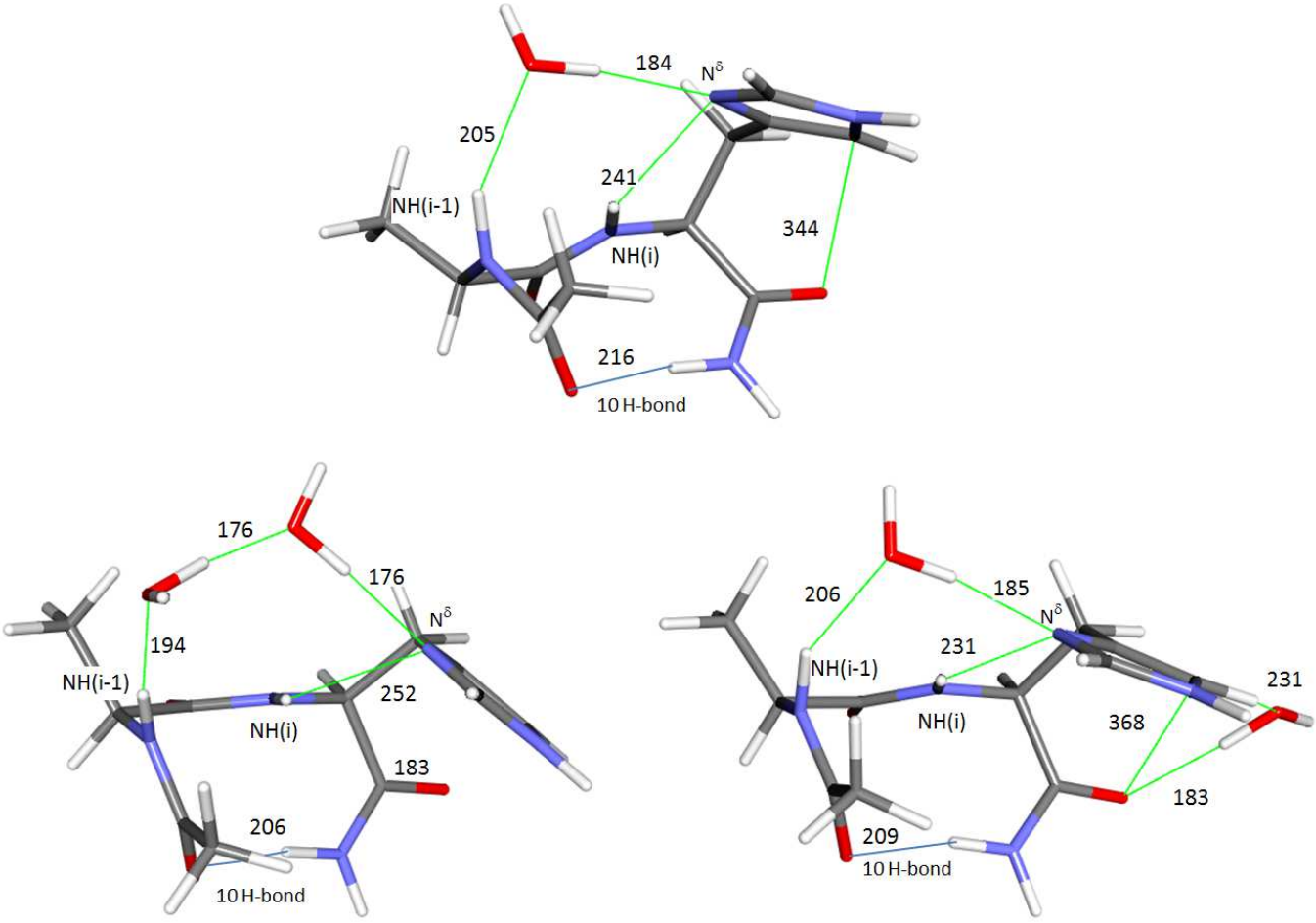

Figure S6: Optimized structure (RI-B97-D3/TZVPP level of theory) of the most stable monohydrate (top) and of two di-hydrates of the ε-6$^{\delta}$ His conformation of Ac-Ala-His-NH_2 in a type I β-turn backbone conformation. One can notice that monohydrate configurations where the water molecule is inserted between the NH(i) and N$^{\delta}$ sites (breakage of the 6$^{\delta}$ bond) are not minima at the quantum chemistry level, even when rotating the His side chain to better accommodate the water molecule; their optimisation converges towards the minimum of Fig 9 left.

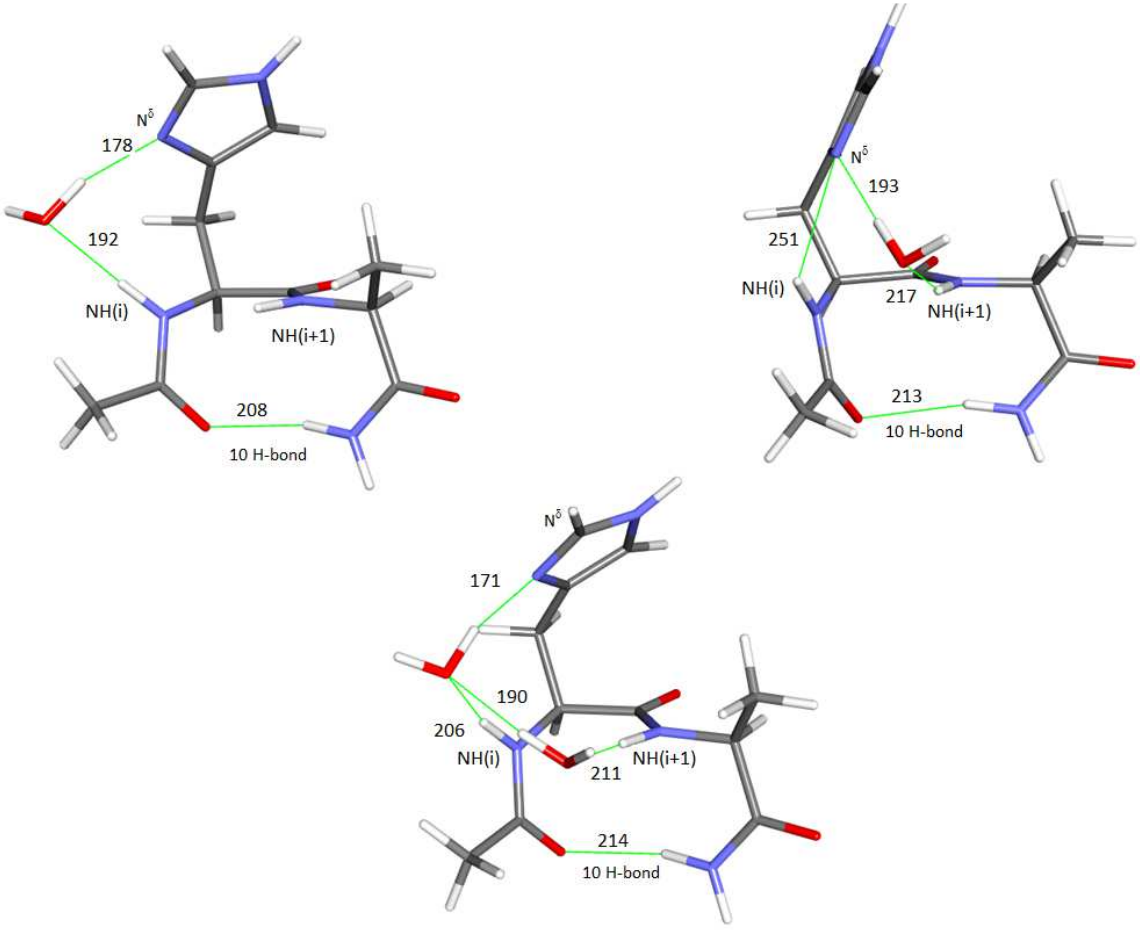

Figure S7: Optimized structure (RI-B97-D3/TZVPP level of theory) of most stable monohydrates (top) and the most stable dihydrate of the ε-6^{δ} His conformation of the Ac-His-Ala-NH_2 model molecule, in a type I β-turn backbone conformation. In the most stable monohydrate (top left), a breakage of the 6^{δ} bond occurs, and a water bridge between the NH(i) and N^{δ} sites takes over. In the next stable form (7 and 14 kJ/mol higher in energy at 0 K and 300 K; top right), the water bridge occurs between the NH(i+1) and N^{δ} sites, leaving the 6^{δ} bond unperturbed. In the dihydrate, when both waters are located on the same side of the His SC, the most stable form corresponds to the hydration of the most stable monohydrate, in which the second water molecule bridges the first one to the NH(i) site.

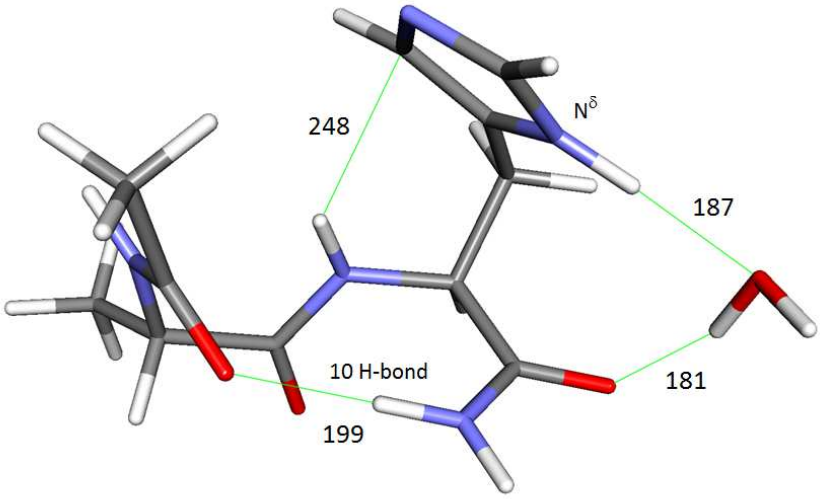

Figure S8: Optimized structure (RI-B97-D3/TZVPP level of theory) of the first monohydrate of the δ-$^{\delta}7/\pi^{H}$ His conformation of Ac-Ala-His-NH_2 in a type I β-turn backbone conformation.

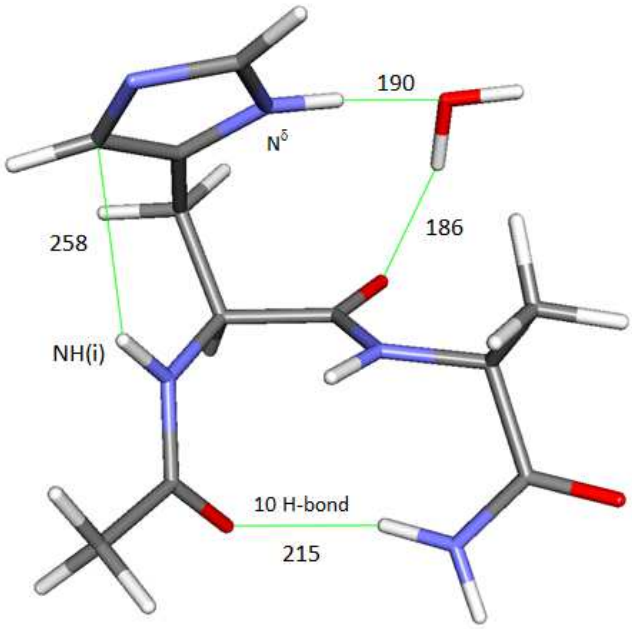

Figure S9 : Optimized structure (RI-B97-D3/TZVPP level of theory) of the first monohydrate of the δ-$^{\delta}7/\pi^{H}$ His conformation of the Ac-His-Ala-NH_2 model molecules in a type I β-turn backbone conformation.

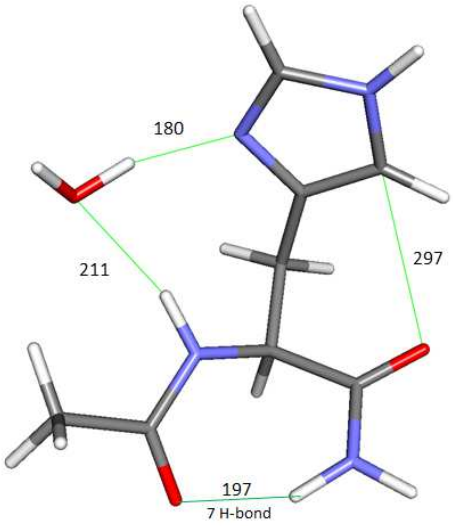

Figure S10 : Optimized structure (RI-B97-D3/TZVPP level of theory) of the most stable monohydrate of the ε-6^{δ} His conformation of the Ac-His-NH_2 model in a γ-turn backbone conformation.

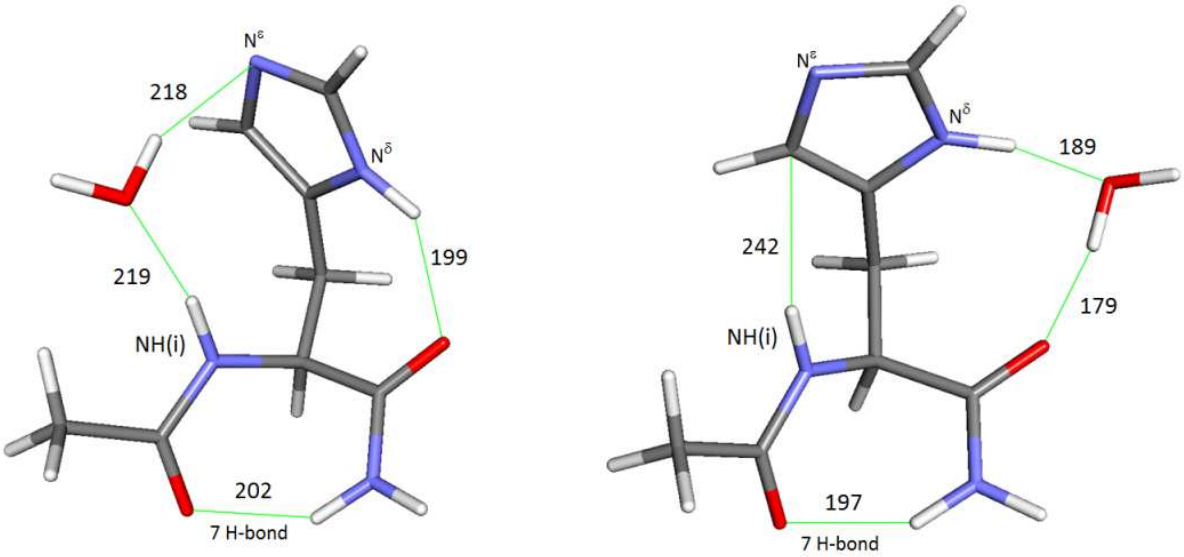

Figure S11: Optimized structure (RI-B97-D3/TZVPP level of theory) of the two most stable monohydrates of the δ-$^{\delta}7/\pi^{H}$ His conformation of the Ac-His-NH_2 model in a γ-turn backbone conformation.

References

1 E. Gloaguen and M. Mons, *Topics in Current Chemistry*, 2015, **364**, 225-270.

2 E. Gloaguen, Y. Loquais, J. A. Thomas, D. W. Pratt and M. Mons, *J. Phys. Chem. B*, 2013, **117**, 4945-4955.