

Nuclear Magnetic Resonance Studies of Thyrotropin-Releasing Hormone (TRH) and Analogues Incorporating D-Histidine and 4-Hydroxy-L-proline

Dagfinn W. Aksnes,^a Constantinos Athanassopoulos,^b Vassiliki Magafa,^b Atle Aaberg,^a George W. Francis,^a and Dionissios Papaioannou^{b,*}

^aDepartment of Chemistry, University of Bergen, Allegaten 41, N-5007 Bergen, Norway and ^bDepartment of Chemistry, School of Natural Sciences, University of Patras, 26500 Patras, Greece.

Aksnes, D. W., Athanassopoulos, C., Magafa, V., Aaberg, A., Francis, G. W. and Papaioannou, D. 1996. NMR Studies of TRH and Analogues Incorporating D-Histidine and 4-Hydroxy-L-proline. – Acta Chem. Scand. 50: 411–416. © Acta Chemica Scandinavica 1996.

NMR studies have been used to examine conformational effects in thyrotropin-releasing hormone (TRH), the epimer incorporating D-His, and their analogues where *trans*- and *cis*-4-hydroxy-L-proline replace L-proline (Pro). In all six compounds the observed overall conformation of the major conformer around the Pro–His amide bond, and the observed increase of the *cis/trans* ratio between the conformers when L-His is replaced by D-His, can be accommodated by assuming that a ten-membered ring is formed by hydrogen bonding between the N–H of the Pro carboxamide function and the N^π-atom of the His imidazole nucleus.

We have recently described an efficient synthesis of the thyrotropin-releasing hormone (TRH, **1**), the epimeric [D-His²]-TRH (**2**), and their analogues (**3–6**) where *trans*- and *cis*-4-hydroxy-L-proline (Hyp and cHyp) replace L-proline (Pro).¹ ¹H NMR studies of the fully protected tripeptides **7–12** showed that the *cis/trans* ratios between the conformers around the His–Pro amide bond are affected both by replacement of L-His by D-His, and by the presence of the hydroxyl group in position 4 of the proline residue. The H(H)-6 protons [parenthesised letters are used here, as elsewhere, to indicate the residue, (G) pyroglutamic acid, (H) histidine and (P) proline or 4-hydroxyproline] appeared as doublets in all three fully protected tripeptides (**7**, **9**, **11**) incorporating L-His, whereas they appeared as doublets of doublets in the corresponding D-His-containing compounds (**8**, **10**, **12**). This was taken to indicate that the His side-chain rotates freely in the L-His derivatives, and this reduces the difference between the two H(H)-6 protons to the extent that they show accidental equivalence. In contrast, the side-chain is locked in the D-His derivatives by an intramolecular H-bond between the His N^π-atom and the Glp–His amide bond, and the shifts remain clearly separated. Examination of Orbit Molecular Building System (OMBS, Cochranes of Oxford, 1972) models indicated that in the L-His derivatives (**7**, **9**, **11**) the *cis* conformation of the His–Pro amide bond, involving an

H-bond between the carbonyl function of this bond and the Glp–His amide bond, predominates: on the other hand in the case of the D-His derivatives (**8**, **10**, **12**) the major conformer has the *trans* arrangement. The presence of the two bulky trityl groups in all of these derivatives (**7–12**) does of course play an important role in the proposed conformations. The structures of all compounds studied in this work may be found in Fig. 1.

NMR studies of the TRH analogues, unprotected free acids, Glp–His–Hyp–OH (**13**) and Glp–His–cHyp–OH (**14**), indicated that the configuration of the 4-OH group in the Pro ring plays a significant role in the determination of the overall shape of the molecules.² Thus in the Hyp derivative (**13**) two H-bonds are evoked, and these hold the His–Pro amide bond in the *cis* conformation. These H-bonds are between the His N^π-H and the carbonyl group of the carboxy function (a ten-membered ring), and between the O–H in position 4 of Pro and the carbonyl group of the Glp–His amide bond (a ten-membered ring). In the case of the cHyp compound (**14**) the His–Pro amide bond adopts the *trans* conformation, and the two hydrogen bonds evoked involve the O–H of the carboxy function with the carbonyl group of the His–Pro amide bond (a seven-membered ring) and the N–H part of the Glp–His amide bond with the O atom of the 4-OH group (an eight-membered ring).

Given the above background, it was of interest to test these models against the unprotected TRH molecule (**1**)

* To whom correspondence should be addressed.

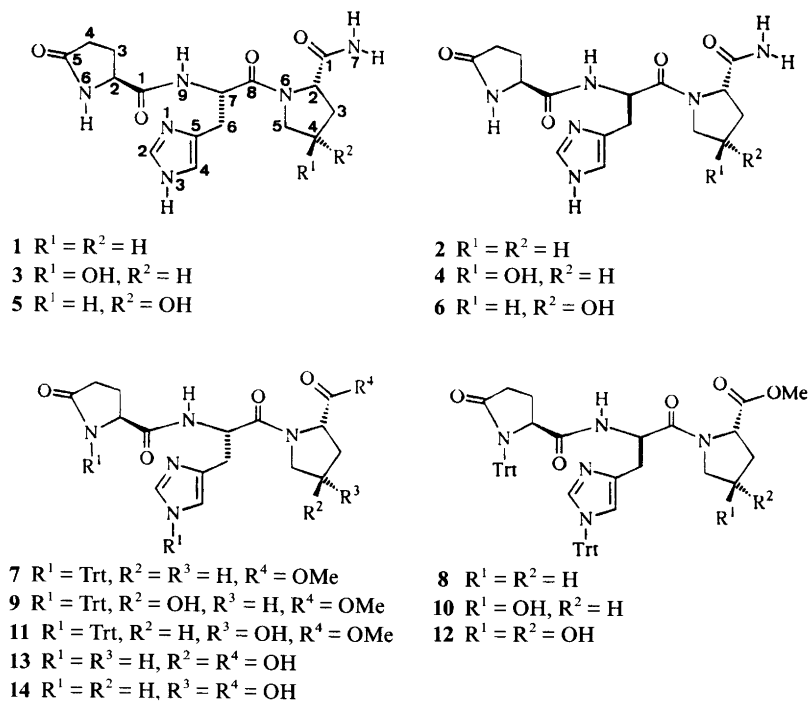


Fig. 1. Structures of the compounds encountered in this work.

and the analogues (2–6), to identify the preferences in the overall shape of TRH caused by the difference between D- and L-His residues, and by the incorporation and stereochemistry of a 4-OH in the Pro residue.

The high-field NMR data for the six TRH related tripeptides are to be found in Table 1, the L-His compounds (1, 3, 5), and Table 2, the D-His compounds (2, 4, 6). A partial 2D NMR spectrum of Glp-L-His-cHyp-NH₂ (5) is reproduced in Fig. 2. Examination of the ¹H NMR spectra of these tripeptides allowed determination of the ratio of the two possible conformers (*cis* and *trans*) about the His-Pro amide bond to be L-Glp-L-His-L-Pro-NH₂ (1), 12:88; L-Glp-L-His-L-Hyp-NH₂ (3), 7:93; L-Glp-L-His-L-cHyp-NH₂ (5), 9:91; L-Glp-D-His-L-Pro-NH₂ (2), 39:61; L-Glp-D-His-L-Hyp-NH₂ (4), 27:73; L-Glp-D-His-L-cHyp-NH₂ (6), 21:79. It should be noted that the *cis/trans* ratios obtained for the pair of dipeptides 1 and 2 are comparable to those previously reported, i.e. 14:86 (in D₂O) and 6:94 (in DMSO-d₆) for 1, and ca. 50:50 (in D₂O) for 2.³

It is apparent from the above ratios that replacing the L-His residue with D-His results in an approximately three-fold increase in the proportion of the minor conformer present in the conformational mixture. On the other hand, incorporation of the OH group in position 4 of the Pro ring has only a small effect on the conformational distribution, while the configuration at C-4 of the Hyp/cHyp ring plays no significant role in this distribution.

Conformational information about the L-His and D-His series was obtained from NOE experiments carried out on compounds 5 and 6. Irradiation of proton H(H)-7

of 5 resulted in enhancement of the H(H)-6a/b and H(P)-5a protons, while similar irradiation of proton H(H)-7 of 6 gave strong enhancement of the H(H)-6a/b and H(P)-5a/b protons. Conversely, irradiation of H(H)-6a of 5 produced a strong enhancement of H(H)-7 and a somewhat weaker enhancement of H(H)-4.

In contrast to the situation for 7–14, the H(H)-6 protons in all six peptides 1–6 appear as a pair of doublets in the regions 3.032–3.120 and 2.909–2.986 ppm, although in the case of 4 these protons give the appearance of broad multiplets. The chemical non-equivalence of the H(H)-6 protons is expected, since the atoms are attached to a prochiral centre. Rapid rotation around the C–C bonds will not make them equivalent but might reduce the chemical shift difference to the extent that the resonances coincide accidentally as observed for the fully protected tripeptides incorporating L-His (7, 9, 11). A broad peak is also observed around 12 ppm in compounds 1–6, suggesting strong H-bonding. These results when taken together suggest that in all of the tripeptides 1–6 the His side-chain is conformationally restricted by a hydrogen bond involving the N^π position of the imidazole ring. As in all previously investigated compounds, the Glp protons in all six tripeptides have similar chemical shift values, multiplicities and coupling constants, and this leads to the assumption that the Glp ring is able to rotate freely in the molecule without interference from any H-bond with other amino acids.

A model based on NMR observations has already been proposed for the restricted rotation of the His side-chain.⁴ The model involves two hydrogen bonds, one between the N^π atom and the N-H part of the Glp-His

Table 1. High-field NMR data for TRH and analogs incorporating 4-hydroxy-L-proline.

AA	Proton	δ -Value (coupling)	Carbon	δ -Value
Glp-L-His-Pro-NH₂ (1)				
Glp	H(G)-2	4.212 (dd, <i>J</i> 8.61 and 3.94 Hz)	C(G)-2	55.37
	H(G)-3/4 ^a	2.50–1.90 (m)	C(G)-3	25.16
			C(G)-4	29.23
	H(G)-6	7.941 (s)	C(G)-1/5 ^b	173.91
				172.09
				170.20
His	H(H)-2	7.677 (s)	C(H)-2	135.03
	H(H)-3	n.o. ^f		
	H(H)-4	7.110 (br.s)	C(H)-4	113.89
	H(H)-6a	3.120 (dd, <i>J</i> 14.48 and 7.48 Hz)	C(H)-5	136.05
	H(H)-6b	2.986 (dd, <i>J</i> 14.48 and 5.50 Hz)	C(H)-6	30.43
	H(H)-7	4.788 (q, <i>J</i> 7.12 Hz)	C(H)-7	51.33
	H(H)-9	8.306 (d, <i>J</i> 7.35 Hz)	C(H)-8 ^b	
Pro	H(P)-2	4.385 (dd, <i>J</i> 7.56 and 3.10 Hz)	C(P)-2	60.20
	H(P)-3/4 ^a		C(P)-3	29.38
			C(P)-4	24.33
			C(P)-5	46.81
	H(P)-5a	3.732 (q, <i>J</i> 8.62 Hz)	C(P)-1	177.59
	H(P)-5b	3.389 (unresolved m)		
	H(P)-7a	8.464 (s)		
	H(P)-7b	7.081 (s)		
Glp-L-His-Hyp-NH₂ (3)				
Glp	H(G)-2	4.215 (dd, <i>J</i> 8.60 and 3.90 Hz)	C(G)-2	55.42
	H(G)-3b	2.042 (unresolved m)	C(G)-3	25.23
	H(G)-3a/4 ^c	2.45–2.12 (m)	C(G)-4	29.19
	H(G)-6	7.971 (s)	C(G)-1/5 ^b	173.83
			172.03	
			170.22	
His	H(H)-2	7.676 (s)	C(H)-2	134.74
	H(H)-3	n.o. ^f		
	H(H)-4	7.109 (br.s)	C(H)-4	113.85
	H(H)-6a	3.087 (dd, <i>J</i> 13.20 and 6.90 Hz)	C(H)-5	135.94
	H(H)-6b	2.954 (dd, <i>J</i> 13.20 and 6.90 Hz)	C(H)-6	30.38
	H(H)-7	4.804 (q, <i>J</i> 6.86 Hz)	C(H)-7	51.22
	H(H)-9	8.284 (d, <i>J</i> 6.86 Hz)	C(H)-8 ^b	
Hyp	H(P)-2/4	4.57–4.33 (m)	C(P)-2	59.10
	H(P)-3 ^c		C(P)-3	37.99
			C(P)-4	68.55
			C(P)-5	54.83
	H(P)-5	3.70–3.52 (unresolved m)	C(P)-1	177.62
	H(P)-7a	8.275 (s)		
		H(P)-7b	7.020 (s)	
Glp-L-His-cHyp-NH₂ (5)				
Glp	H(G)-2	4.205 (dd, <i>J</i> 8.84 and 4.20 Hz)	C(G)-2	55.36
	H(G)-3b ^e	2.06–1.90 (br. m)	C(G)-3	25.16
	H(G)-3a/4 ^d	2.46–2.14 (m)	C(G)-4	29.20
	H(G)-6	7.959 (s)	C(G)-1/5 ^b	174.13
			172.14	
			170.41	
His	H(H)-2	7.661 (s)	C(H)-2	135.50
	H(H)-3	n.o. ^f		
	H(H)-4	7.092 (br.m)	C(H)-4	113.69
	H(H)-6a	3.100 (dd, <i>J</i> 13.93 and 7.50 Hz)	C(H)-5	135.89
	H(H)-6b	2.951 (dd, <i>J</i> 13.93 and 5.06 Hz)	C(H)-6	30.28
	H(H)-7	4.762 (q, <i>J</i> 6.68 Hz)	C(H)-7	51.32
	H(H)-9	8.358 (d, <i>J</i> 6.97 Hz)	C(H)-8 ^b	
cHyp	H(P)-2/4	4.363 (br. m)	C(P)-2	59.02
	H(P)-3a ^d		C(P)-3	37.32
	H(P)-3b	2.06–1.90 (m)	C(P)-4	69.94
	H(P)-5a	3.963 (dd, <i>J</i> 9.87 and 5.16 Hz)	C(P)-5	54.88
	H(P)-5b	3.235 (br. d, <i>J</i> 9.87 Hz)	C(P)-1	177.68
	H(P)-7a	8.229 (s)		
		H(P)-7b	7.001 (s)	

^a In addition to the four H(G)-3/4 protons, the four H(P)-3/4 protons also resonate in the same δ region. ^b Unambiguous assignment of the three carbons C(G)-1/5 and C(H)-8 could not be obtained. ^c In addition to the three H(G)-3a/4 protons, the H(P)-3 protons also resonate in the same δ region. ^d In addition to the three H(G)-3a/4 protons, the H(P)-3a proton also resonates in the same δ region. ^e In addition to the H(G)-3b proton, the H(P)-3b proton also resonates in the same δ region. ^f Not observed.

Table 2. High-field NMR data for [D-His²]-TRH and analogs with 4-hydroxy-L-proline.

AA	Proton	δ -Value (coupling)	Carbon	δ -Value	
Glp-D-His-Pro-NH₂ (2)					
Glp	H(G)-2	4.209 (dd, <i>J</i> 8.59 and 4.21 Hz)	C(G)-2	55.27	
	H(G)-3/4 ^a	2.50–1.90 (m)	C(G)-3	25.34	
			C(G)-4	30.54	
H(G)-6	7.955 (s)	C(G)-1/5 ^b	173.62		
			172.50		
His	H(H)-2	7.681 (s)	C(H)-2	169.75	
	H(H)-3	n.o. ^g		134.69	
	H(H)-4	6.930 (s)	C(H)-4	n.o. ^g	
	H(H)-6a	3.056 (dd, <i>J</i> 14.26 and 7.03 Hz)	C(H)-5	n.o. ^g	
	H(H)-6b	2.928 (dd, <i>J</i> 14.26 and 7.30 Hz)	C(H)-6	31.59	
	H(H)-7	4.883 (q, <i>J</i> 7.32 Hz)	C(H)-7	50.90	
	H(H)-9	8.460 (d, <i>J</i> 7.32 Hz)	C(H)-8 ^b		
	Pro	H(P)-2	4.315 (dd, <i>J</i> 8.02 and 3.07 Hz)	C(P)-2	59.47
		H(P)-3/4 ^a	3.745 (ddd, <i>J</i> 14.28, 4.72 and 2.05 Hz)	C(P)-3	30.47
C(P)-4				23.91	
H(P)-5a	3.745 (ddd, <i>J</i> 14.28, 4.72 and 2.05 Hz)	C(P)-5	46.42		
H(P)-5b	3.417 (dt, <i>J</i> 14.28 and 6.92 Hz)	C(P)-1	177.58		
H(P)-7	7.200 (s) and 7.056 (s)				
Glp-D-His-Hyp-NH₂ (4)					
Glp	H(G)-2	4.170 (dd, <i>J</i> 8.35 and 3.89 Hz)	C(G)-2	55.48	
	H(G)-3/4 ^c	2.40–1.90 (m)	C(G)-3	25.54	
			C(G)-4	29.20	
H(G)-6	7.275 (br. s)	C(G)-1/5 ^b	177.48		
			173.59		
His	H(H)-2	7.660 (s)	C(H)-2	172.21	
	H(H)-3	n.o. ^g		134.50	
	H(H)-4	7.381 (s)	C(H)-4	113.20	
	H(H)-6a	3.100 (br. m)	C(H)-5	134.81	
	H(H)-6b	2.909 (br. m)	C(H)-6	39.95	
	H(H)-7	4.863 (br. m)	C(H)-7	58.65	
	H(H)-9	8.450 (br. m)	C(H)-8 ^b		
	Hyp	H(P)-2/4	4.50–4.30 (m)	C(P)-2	58.70
		H(P)-3 ^c	3.85–3.55 (m)	C(P)-3	37.82
C(P)-4				68.57	
H(P)-5	3.85–3.55 (m)	C(P)-5	54.80		
H(P)-7a	6.971 (br. s) ^d	C(P)-1	177.51		
H(P)-7b	6.890 or 7.381 (br. s) ^d				
Glp-L-His-cHyp-NH₂ (6)					
Glp	H(G)-2	4.321 (dd, <i>J</i> 9.15 and 4.01 Hz)	C(G)-2	58.79	
	H(G)-3b ^e	1.97–1.85 (m)	C(G)-3	25.62	
	H(G)-3a/4 ^f	2.43–2.10 (m)	C(G)-4	29.22	
			C(G)-1/5 ^b	174.11	
H(G)-6	7.955 (s)	C(G)-1/5 ^b	172.50		
			170.09		
His	H(H)-2	7.678 (s)	C(H)-2	135.00	
	H(H)-3	n.o.			
	H(H)-4	6.930 (br. s)	C(H)-4	n.o. ^g	
	H(H)-6a	3.032 (dd, <i>J</i> 14.23 and 6.95 Hz)	C(H)-5	n.o. ^g	
	H(H)-6b	2.911 (dd, <i>J</i> 14.23 and 7.02 Hz)	C(H)-6	29.15	
	H(H)-7	4.385 (q, <i>J</i> 7.29 Hz)	C(H)-7	51.14	
	H(H)-9	8.492 (d, <i>J</i> 7.41 Hz)	C(H)-8 ^b		
	cHyp	H(P)-2/4	4.27–4.16 (m)	C(P)-2	55.45
		H(P)-3a ^f	C(P)-3	37.12	
H(P)-3b ^e		C(P)-4	69.00		
H(P)-5a	3.614 (dd, <i>J</i> 10.47 and 5.15 Hz)	C(P)-5	54.92		
H(P)-5b	3.555 (dd, <i>J</i> 10.47 and 3.66 Hz)	C(P)-1	177.50		
H(P)-7a	7.280 (s)				
H(P)-7b	7.171 (s)				

^a In addition to the four H(G)-3/4 protons, the four H(P)-3/4 protons also resonate in the same δ region. ^b Unambiguous assignment of these two carbons and carbon C(H)-8 could not be obtained. ^c In addition to the four H(G)-3/4 protons, the H(P)-3 protons also resonate in the same δ region. ^d Not unambiguously identified. ^e In addition to the H(G)-3b proton, the H(P)-3b proton also resonates in the same δ region. ^f In addition to the three H(G)-3a/4 protons, the H(P)-3a proton also resonates in the same δ region. ^g Not observed.

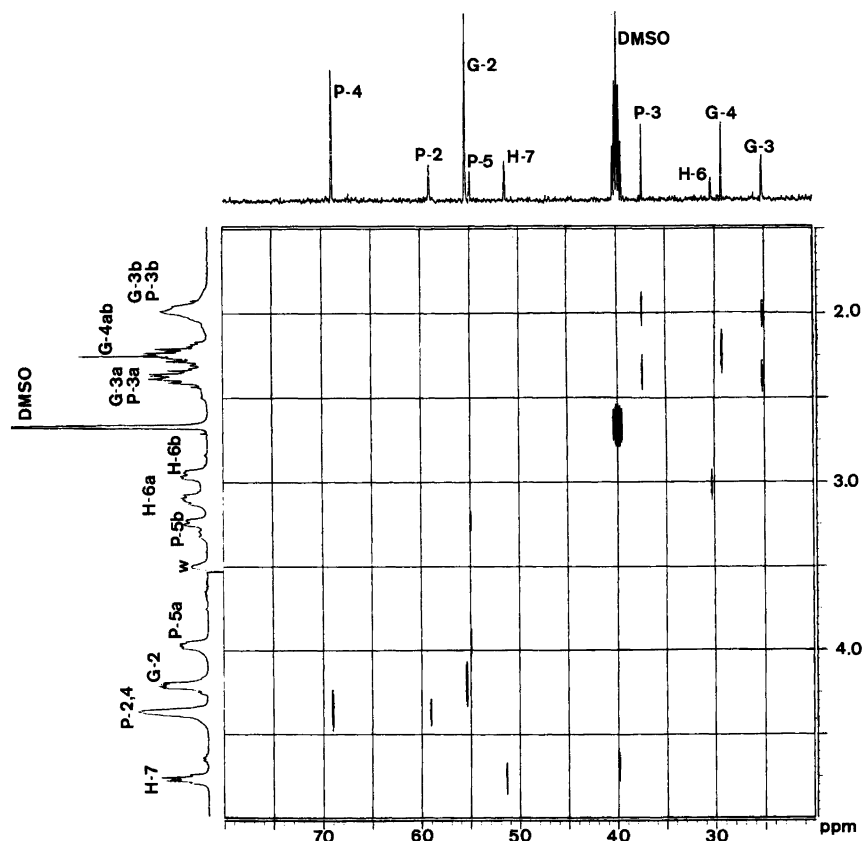


Fig. 2. Partial 2D NMR spectrum of Glp-L-His-cHyp-NH₂ (5) in DMSO. A weak signal due to the presence of water is marked with the letter W.

amide bond and the other between the *N*-H portion of the prolinamide function and the carbonyl part of the Pro-His bond. This model requires that the His-Pro amide bond adopts the *trans* arrangement. An alternative model has been proposed for the structurally similar tripeptide L-Glp-L-His-Gly and involves a ten-membered ring formed by a hydrogen bond between the *N*^π atom and the OH part of the glycine (Gly) carboxy group.⁵ This last model is quite similar to the one previously proposed by us for the tripeptide acid 13.²

An examination of the OMBS models for compounds 1–6 revealed that a ten-membered ring forming hydrogen bond can be nicely accommodated (a linear *N*-H-*N*^π bond) in all six compounds, taking into consideration the configuration of the His asymmetric center and the *cis/trans* conformational isomerisation about the His-Pro amide bond. Furthermore, examination of the models for the two conformers about the His-Pro amide bond revealed that the *trans* conformation can readily accept a change in the configuration of the central amino acid (His). The *cis* configuration can also accommodate within a ten-membered ring both L- and D-His residues, although in the case of the L-configuration unfavourable steric congestion may exist between the Glp-NH residue and the Pro ring. The observed *cis/trans* ratios may thus be taken to suggest that the major conformer about the

His-Pro amide bond is of the *trans* type in both L- and D-series.

Additional evidence was obtained from the NOE experiments mentioned above. These experiments showed that there is strong, through-space interaction between proton H(H)-7 and H(P)-5a in the TRH-L, and proton H(H)-7 and H(P)-5a/b in one form of TRH-D. These interactions, which imply that the protons involved are fairly close in space, can only be explained if the Pro-His amide bond adopts the *trans* conformation. The proposed models for the major conformers, TRH-L and TRH-D respectively, for the L- and D-series are drawn in Fig. 3. Both models show that the hydroxy group in position 4 of the Pro ring can not play a significant role in deciding the overall shape of the molecule, e.g. by forming a second hydrogen bond with the remaining Glp-NH portion of the molecule. In addition, these two models can be of either subtype A or B. These subtypes are realised by rotation around the σ -bond between C(H)-6 and C(H)-7. In the case of the L-series, both subtypes involve through-space interaction between protons H(H)-7 and H(P)-5a: such interaction was indeed observed. In the case of subtype A, however, an additional interaction not available in subtype B could be expected between protons H(H)-6a and H(P)-5b. Unfortunately, the H(P)-5b resonance was too close to the irradiation point

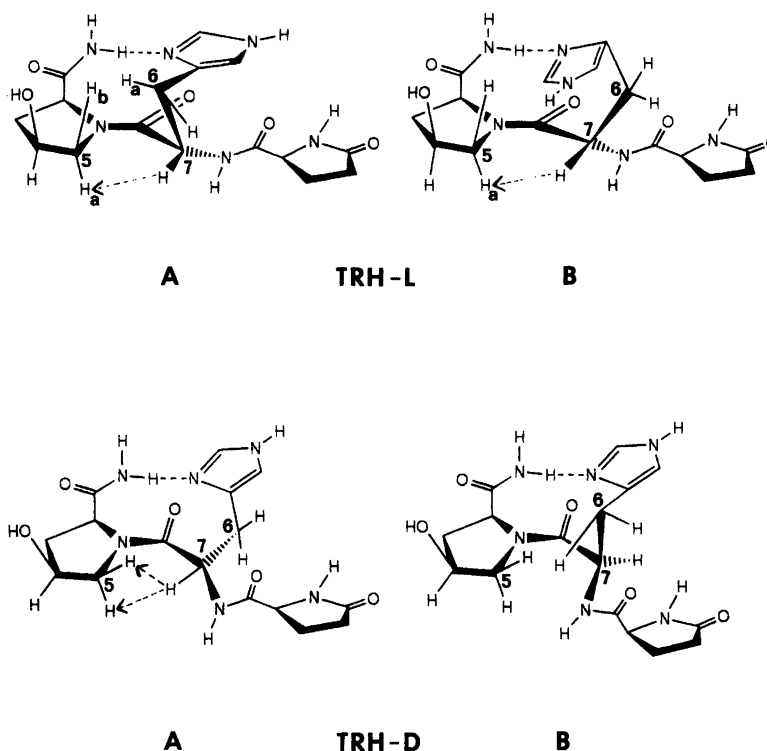


Fig. 3. Proposed models for TRH and analogs incorporating L-His (TRH-L) or D-His (TRH-D) and 4-hydroxy-L-proline.

to see any NOE, and thus this uncertainty remains. On the other hand in the case of the D-series, a strong through-space interaction, only possible in subtype A of model TRH-D, was observed between protons H(H)-7 and H(P)-5a/b.

Biological evaluation of the six tripeptides, and synthesis and proton NMR studies of various *N*-methylated TRH and [*D*-His²]-TRH analogs, are now in progress.

Experimental

Compounds 1–6. The six TRH-related tripeptides were obtained by solution peptide synthesis using as key intermediates the dipeptides Trt-Glp-L- and D-His(Trt)-OMe.¹ These dipeptides are readily available in high yields from the coupling of the 1-hydroxybenzotriazolyl ester of *N*-tritylpyroglutamic acid (Trt-Glp-OBt) with *N*^{im}-Trt-L- and D-His-OMe, respectively.

NMR Measurements. All compounds were dissolved in dimethyl sulfoxide (DMSO-*d*₆) in 5 mm o.d. sample tubes: D₂O was not used to avoid H-exchange and problems with broad-band viscosity. The 400.13 and 100.62 MHz ¹H and ¹³C NMR spectra, respectively, were obtained at 298 K on a Bruker AM 400 WB spectrometer. The deuteromethyl carbon signal and the residual proton signal of the solvent were used as secondary references for the chemical shifts (39.50 and 2.62 ppm,

respectively). All the 90° transmitter and decoupler pulses were carefully calibrated (6.7–14.5 μs).

The ¹H NMR spectra were obtained with presaturation of the solvent signal. The homonuclear Overhauser enhancements were determined by subtracting the unperturbed FID with off-resonance irradiation from the perturbed FID with on-resonance irradiation followed by Fourier transformation and phasing. The broad-band decoupled ¹³C NMR spectra were acquired using the standard one-pulse, DEPT and spin-echo experiments.

The ¹H-¹H and ¹H-¹³C chemical shift correlations were established using the standard COSY and HSC shift correlation pulse sequences. The heteronuclear 2D, spin-echo and DEPT experiments were optimised for one-bond couplings of 140 Hz.

References

1. Papaioannou, D., Athanassopoulos, C., Magafa, V., Karigiannis, G., Karamanos, N., Stavropoulos, G., Napoli, A., Sindona, G., Aksnes, D. W., Francis, G. W. and Aaberg, A. *Acta Chem. Scand.* 49 (1995) 103.
2. Stavropoulos, G., Karagiannis, K., Vynios, D., Papaioannou, D., Aksnes, D. W., Frøystein, N. Å. and Francis, G. W. and Aaberg, A. *Acta Chem. Scand.* 45 (1991) 1047.
3. Deslauriers, R., Garrigou-Lagrange, C., Bellocq, A.-M. and Smith, I. C. P. *FEBS Lett.* 31 (1973) 59.
4. Fermandjian, S., Pradelles, P. and Fromageot, P. *FEBS Lett.* 28 (1972) 156.
5. Gorbitz, C. H. and Krane, J. *Acta Chem. Scand.* 47 (1993) 979.

Received April 7, 1995.