Supplementary Information

for

The reversible macrocyclization of Tyrocidine A aldehyde: a hemiaminal reminiscent of the tetrahedral intermediate of macrolactamization

Sebastian Enck, Florian Kopp, Mohamed A. Marahiel* and Armin Geyer* geyer@staff.uni-marburg.de

Peptide synthesis and purification	p. S1
Characterization by HR-ESI-MS	p. S3
NMR measurements and data	p. S7
pH and temperature dependence of TycA-CHO macrocyclization equilibrium	p. S12
Estimation of the macrocyclization entropy	p. S16
NOE-based molecular modeling of Ψ [<i>R</i> -CH(OH)NH]-TycA	p. S17
References	p. S17

Peptide synthesis and purification

The linear cyclization precursor TycA-CHO (Scheme S1) was synthesized on solid phase according to ref^{S1}. NovaSyn[®] TG resin (NovaBiochem) with the C-terminal leucine aldehyde immobilized as oxazolidine was used. The resin was left swelling in DMF for 30 min and was then suitable for automated Fmoc solid phase synthesis carried out on an Advanced ChemTech APEX 396 synthesizer (0.1 mmol scale). After the last peptide coupling, protecting groups were removed with 100% TFA (2 x 10 min), the resin was washed with DCM, and the peptide aldehyde was cleaved from resin by treatment with DCM/MeOH/AcOH/H₂O (63:22:10:5) for 30 min. The crude peptide aldehyde was precipitated in hexane, dissolved in acetonitrile/H₂O and purified by preparative HPLC on an Agilent 1100 system with a reversed-phase Nucleodur C18 column (Macherey-Nagel, 250/21, particle size 100 Å, pore diameter 3 μ m). A gradient 5 - 55% of 0.1% TFA/acetonitrile in 0.1% TFA/H₂O was applied over 30 min (20 ml/min, 25 °C). After lyophilization, TycA-CHO was obtained as a colourless solid (9.1 mg, yield 7.2 % based on resin loading).



Scheme S1 Overview of TycA-CHO cyclization equilibrium. In aqueous solution at pH 3.0 the aldehyde hydrate TycA-CH(OH)₂ (yellow) is predominantly present (< 10% aldehyde (orange)). Upon macrocyclization, one aminal epimer is selectively formed (Ψ [CH(OH)NH]-TycA, green), and no dehydration which would give the imine (Ψ [C=N]-TycA, red) is observed.

Characterization by HR-ESI-MS

The identity of the peptide aldehyde obtained after preparative HPLC and lyophilization was confirmed by analytical HPLC-ESI-MS on an Agilent 1100 system with a reversed-phase Nucleodur 100-5 C18 column (Macherey-Nagel, 125/2, particle size 100 Å, pore diameter $3 \mu m$). A gradient 10 - 95% of 0.05% HCOOH/acetonitrile in 0.05% HCOOH/H₂O was applied over 50 min (0.2 ml/min, 45 °C). Fig. S1 shows the UV trace of the HPLC-MS analysis (a) as well as the FTMS ion chromatograms for the masses corresponding to TycA-CH(OH)₂ (b) and TycA-CHO (c). Recalling the fast equilibria between the species, the conditions of ESI-MS analysis can not be compared to the conditions used for NMR spectroscopic analysis. The integrated mass spectrum over the main peak retention time (20.78 - 22.03 min) is shown in Fig. S2a and the section of m/z = 630 - 660 is shown in Fig. S2b. In Fig. S2c and d, the calculated high-resolution ESI spectra for the linear species are depicted.



Fig. S1 LC-ESI-MS analysis of TycA-CHO. (a) UV trace (215 nm). (b) FTMS spectrum for m/z = 1290.0 - 1291.0 (Tyc-CH(OH)₂). (c) FTMS spectrum for m/z = 1272.0 - 1273.0 (TycA-CHO).



Fig. S2 Integrated mass spectra (retention time: 20.78 - 22.03 min; Fig. S1) of HPLC-ESI-MS analysis of TycA-CHO under acidic conditions. (a) Overview of ESI spectrum. (b) high-resolution ESI spectrum of area m/z = 630 - 660. (c) Calculated high-resolution ESI spectrum of $[TycA-CH(OH)_2 + 2H]^{2+}$. (d) Calculated high-resolution ESI spectrum of $[TycA-CHO + 2H]^{2+}$. The asterisks * mark signals which are caused by small amounts (< 10%) of dimers.

The macrocyclic hemiaminal Ψ [CH(OH)NH]-TycA and the linear aldehyde Tyc-CHO have molecular weigth $([M + 2H]^{2+} = 636.84)$. As a cyclic methyl identical aminal $(\Psi[CH(OMe)NH]-TycA)$ in contrast would be identifiable by its $[M + 2H]^{2+} = 643.85$ signal we performed an additional experiment. A small amount of basic aqueous solution was evaporated and the residue was subsequently dissolved in MeOH. The residue obtained after evaporation of the MeOH was finally dissolved in basic aqueous solution at pH 9 and subjected to ESI-MS analysis. In fact, a weak methyl hemiaminal signal could be identified (Fig. S3). Besides a further weak signal which originates from the cyclic hemiaminal and/or the linear aldehyde, the main signal represents the cyclic imine Ψ [C=N]-TycA. The fact that small amounts of Ψ [CH(OMe)NH]-TycA are still observed speaks for its stability which fits well of the great inclination of TycA-CHO to form a hemiaminal in H₂O. The full ESI-MS spectrum is shown in Fig. S3a, and the section of m/z = 598 - 702 is shown in Fig. S3b. The calculated high-resolution ESI spectra of the cyclic species identified are given in Fig. S3c, d, and e. In Table S1 all calculated and observed masses are listed.

Table S1	Calculated and observed m	asses from HR-ESI-MS	; (in gmol⁻¹) c	of linear and cyclic pep	tides.
All masse	s refer to the [M + 2H] ²⁺ peal	ks.			

compound	molecular formula	calculated mass	observed mass
TycA-CHO	$C_{66}H_{89}N_{13}O_{13}$	636.8424	636.8437
TycA-CH(OH) ₂	$C_{66}H_{91}N_{13}O_{14}$	645.8477	645.8488
Ψ[C=N]-TycA	$C_{66}H_{87}N_{13}O_{12}$	627.8371	627.8366
Ψ[CH(OH)NH]-TycA	$C_{66}H_{89}N_{13}O_{13}$	636.8424	636.8430
Ψ[CH(OMe)NH]-TycA	$C_{67}H_{91}N_{13}O_{13}$	643.8502	643.8507

S6



Fig. S3 HR-ESI-MS analysis of a basic peptide solution after treatment with MeOH. (a) Overview of ESI spectrum. (b) High-resolution ESI spectrum of area m/z = 598 - 702. (c) Calculated high-resolution ESI spectrum of $[\Psi[TycA-CH(OH)NH] + 2H]^{2+}$ (and $[TycA-CHO+ 2H]^{2+}$). (d) Calculated high-resolution ESI spectrum of $[\Psi[TycA-C=N] + 2H]^{2+}$. (e) Calculated high-resolution ESI spectrum of $[\Psi[TycA-C=N] + 2H]^{2+}$. CH(OMe)NH] + 2H]²⁺.

NMR measurements and data

All measurements were performed on a Bruker Avance DRX 600 spectrometer with a 5 mm BBI probe head. The peptide concentrations in the samples used were 2.35, 3.82, and 4.68 mmol/l, and the SDS- d_{25} concentration was 0.19 mmol/l (see p. S12 for details). Water suppression was achieved by excitation sculpting with gradients (double watergate DPFGSE sequence).^{S2} Homonuclear 2D spectra (TOCSY, NOESY) were recorded in the phasesensitive mode as data matrices of 512 (t_1) real x 2048 (t_2) complex data points; 16 or 32 scans were used per t_1 increment. The used spectral widths were 6010 Hz in each dimension for both TOCSY and NOESY experiments. Mixing times of 100 ms (TOCSY) and 150 to 200 ms (ROESY) were applied. Heteronuclear 2D HSQC experiments were performed in the phase-sensitive mode with data matrices of 512 or 1024 (t_1 , ¹³C) real x 2048 (t_2 , ¹H) complex data points and 16 to 32 scans per t_1 increment. The used spectral widths were 6010 (¹H) / 22640 Hz (¹³C). All data were recorded and analyzed using Bruker TopSpin software. The spectra of the samples containing $SDS-d_{25}$ were calibrated on the residual SDS methyl signal which appears temperature independent at 0.772 ppm.^{S3} The spectra obtained from the sample without SDS- d_{25} were calibrated on the polyethylene glycol (PEG) signal (residual PEG from resin) which occurs at 3.637 ppm in the spectra calibrated on the residual SDS methyl signal. The assigned ¹H NMR spectra are depicted in Fig. S4 and S5, and all proton chemical shifts are listed in Table S2.

The temperature dependence of the macrocyclization equilibrium (Fig. S7 and S8) as well as the temperature gradients (Table S3) were analyzed by recording ¹H NMR spectra of the linear/cyclic mixture at pH 6.6 between 290 and 330 K with 10 K increments. The NMR sample was allowed to equilibrate in the spectrometer at the respective temperature for 1 h prior to the measurements.

Table S2 ¹H chemical shifts (in ppm) of TycA species (600 MHz, 300 K) in KH_2PO_4/H_3PO_4 buffer (H_2O/D_2O 5:1).

Left: TycA-CH(OH)₂ without SDS- d_{25} at pH 3.0. Middle: TycA-CH(OH)₂ in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 3.0.

Right: Ψ [CH(OH)NH]-TycA in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 9.0.

Desitio	2	TycA-CH(OH) ₂	TycA-CH(OH) ₂	Ψ[CH(OH)NH]-TycA
POSITIO	n		SDS- <i>d</i> ₂₅	SDS-d ₂₅
D-	NH3 ⁺	(a)	(a)	
Phe1	hemiaminal-NH			5 3 3
1 1101	C -H	4 46	4 49	3 73
	C₀-H	3.07	3 13	2 67
	CβII	3.17	3 39	2.07
	Cac-H	(b)	(b)	(b)
	C _{2,6} H	(b)	(b)	(b)
	С4-Н	(b)	(b)	(b)
	04 11	(0)	(0)	
Pro2	C _α -H	4.17	4.14	3.99
	C _β -H	1.43	1.24	1.62
		1.81	1.60	1.74
	C _γ -H	1.35	1.04	0.86
			1.15	1.13
	C _δ -Н	2.60	3.39	2.33
		3.35	3.42	3.22
		o		<
Phe ₃	NH	8.12	7.66	6.99
	C_{α} -H	4.54	4.58	4.77
	С _β -Н	2.82	2.82	2.18
	C II	2.95	3.05	2.68
	С _{2,6} -Н	(b)	(b)	(b)
	С _{3,5} -Н	(b)	(b) (1)	(b)
	С4-Н	(0)	(0)	(0)
D-	NH	8.00	7.90	8.68
Phe4	C _a -H	4.43	4.58	5.60
	С _в -Н	2.82	3.11	2.89
	Ч	2.90	3.23	3.07
	C ₂₆ -H	(b)	(b)	(b)
	C _{3.5} -H	(b)	(b)	(b)
	C ₄ -H	(b)	(b)	(b)
Asn5	NH	8.33	8.18	9.00
	C _α -H	4.53	4.49	4.89
	C _β -H	2.49	2.54	3.04
			2.68	3.35
	CONH_2	(a)	(a)	(a)
Gln6	NH	8.18	8.17	8.33
0	C _a -H	4 15	4 27	4 00
	C _B -H	1.82	1.91	1.68
	- p	1.87		1.83
	C _ν -H	2.09	2.12	1.60
	r	2.14		1.82
	CONH ₂	(a)	(a)	(a)

(continued on next page)

Tyr7	NH	8.12	8.20	8.19
-	C _α -H	4.48	4.44	4.50
	C _β -H	2.84	2.88	2.91
	r	2.98	3.07	3.13
	C ₂₆ -H	7.02	7.11	7.05
	C ₃₅ -H	6.73	6.78	6.79
	OH	(a)	(a)	(a)
Val8	NH	7.73	7.73	7.73
	C _a -H	3 95	4 12	4 39
	Ce-H	1 88	2.09	2.10
	β-Me	0.81	0.90	1 01
	p	0.01	0.94	1.01
Orn9	NH	8 25	7 96	8 37
01112	C -H	4 22	4 43	5.16
	Cα-H	1 79	1.13	1 88
	CβII	1.75	1.85	2 04
	C -H	1 66	1.65	1 67
	Ογ Π	1.00	1.07	1 71
	Cs-H	2 96	3.00	2 87
	δ -NH ₂ ⁺	7 53	7 44	(a)
	0 1113	1.55	/	(u)
Leu10	NH	7.84	7.49	7.73
	C _α -H	3.83	3.91	4.19
	C _β -H	1.33	1.37	1.31
				1.34
	C _γ -H	1.45	1.52	1.63
	γ-Me	0.75	0.82	0.97
		0.83	0.85	
	СНО	9.42	9.42	
	C <u>H(</u> OH) ₂	4.86	4.88	
	$CH(OH)_2$	(a)	(a)	
	C <u>H(</u> OH)NH			4.43
	CH(OH)NH			(a)

(continued from previous page)

(a) Precise assignment was not possible due to fast chemical exchange with water.

(b) The Phe aromatic protons could not be precisely assigned due to signal overlap.

Table S3 Temperature gradients (in ppb K⁻¹) of TycA-CH(OH)₂ and Ψ [CH(OH)NH]-TycA in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) in KH₂PO₄/H₃PO₄ buffer (H₂O/D₂O 5:1).

position	TycA-CH(OH) ₂	Ψ [CH(OH)NH]-TycA
D-Phe1 hemiaminal-NH		- 15.0
Phe3-NH	- 4.7	- 4.3
D-Phe4-NH	- 2.5	- 3.2
Asn5-NH	- 3.3	- 4.3
Gln6-NH	(a)	- 3.1
Tyr7-NH	- 4.3	-4.0
Val8-NH	- 3.0	+ 1.0
Orn9-NH	- 4.7	- 5.0
Leu10-NH	- 4.0	+ 1.0

(a) No data could be obtained due to signal overlap.



Fig. S4 Assigned ¹H NMR spectra (watergate, 600 MHz, 300 K, KH_2PO_4/H_3PO_4 buffer, H_2O/D_2O 5:1), up-field range. (a) Ψ [CH(OH)NH]-TycA in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 9.0. (b) TycA-CH(OH)₂ in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 3.0. (c) TycA-CH(OH)₂ without SDS- d_{25} at pH 3.0.



Fig. S5 Assigned ¹H NMR spectra (watergate, 600 MHz, 300 K, KH₂PO₄/H₃PO₄ buffer, H₂O/D₂O 5:1), down-field range. (a) Ψ [CH(OH)NH]-TycA in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 9.0. (b) TycA-CH(OH)₂ in the presence of SDS- d_{25} (80 eq, 0.19 mol/l) at pH 3.0. (c) TycA-CH(OH)₂ without SDS- d_{25} at pH 3.0. The asterisks * mark the Leu10 CHO peaks resulting from the linear aldehyde TycA-CHO which accounts for less than 10% at pH 3.0.

pH and temperature dependence of TycA-CHO macrocyclization equilibrium

The NMR sample without SDS- d_{25} was prepared by dissolving 3.50 mg (2.75 µmol) of TycA-CHO in 0.72 ml of a partially deuterated H₃PO₄/KH₂PO₄ buffer solution (H₂O/D₂O 5:1) at pH 3.0 (TycA-CHO concentration: 3.82 mmol/l). Under these conditions, only the linear peptide could be examined (Fig. S4c and S5c) as precipitation was observed when rising the pH. This likely results from formation of the cyclopeptide which is well-known to aggregate in aqueous solution.^{S4} To enable measurements at higher pH which is necessary for the NMR spectroscopic observation of cyclic species and of cyclization equilibria, perdeuterated sodium dodecylsulfate (SDS- d_{25}) was chosen as membrane mimetic. The NMR sample was prepared by dissolving 1.80 mg (1.41 µmol) of TycA-CHO and 36 mg (82 eq, 115 µmol) of SDS- d_{25} in 0.6 ml of a partially deuterated H₃PO₄/KH₂PO₄ buffer solution (H₂O/D₂O 5:1) at pH 3.0 (TycA-CHO concentration: 2.35 mmol/l, SDS- d_{25} concentration: 0.19 mol/l). Under these conditions, no precipitation was observed in the pH range between 3.0 and 9.0. For some further measurements, additional 1.40 mg (1.10 µmol) of TycA-CHO were added, resulting in a TycA-CHO concentration of 4.68 mmol/l.

The TycA-CHO cyclization equilibrium (Scheme S1) was controlled directly in the NMR tube. The macrocyclization of TycA-CH(OH)₂ in the presence of SDS- d_{25} was achieved by addition of solid Na₂CO₃ until the pH reached 9.0. After approx. 20 min complete cyclization to Ψ [CH(OH)NH]-TycA was observed in the ¹H NMR spectrum.

For the examination of the pH dependence of the cyclization equilibrium, the pH was raised or lowered by addition of solid Na₂CO₃ and aqueous H₃PO₄, respectively. A series of ¹H NMR spectra recorded at different pH values and points of time is shown in Fig. S6, and spectra recorded at different temperatures are shown in Fig. S7 and S8. After every temperature change the system was allowed to equilibrate for 1 h in the spectrometer. All pH and temperature changing experiments were carried out with the same peptide sample in order to prove the reversibility of the system.

As discussed in the main section, a second set of signals becomes visible at high pH which at room temperature slowly grows in intensity over several hours and considerably increases at higher temperature. The corresponding signals are marked by grey bars in Fig. S6 and S7. As this set of signals is only visible if Ψ [CH(OH)NH]-TycA is present (Fig. S6) and as no imine or further linear species is detected, the corresponding peptide most likely is another macrocycle which slowly and reversibly emerges from the hemiaminal main product. Although the formation of cyclic dimers can not be completely excluded, the existence of such species is improbable as it was never observed in the irreversible cyclization of TycA.^{S5}

The variation of temperature has a significant effect on the amount of the secondary product (Fig. S7). As mentioned in the main section, this transformation can either be an epimerization of the hemiaminal stereocenter which is generated upon macrocyclization. This would involve the release and subsequent addition of H₂O via an imine intermediate. The second possibility, which is the epimerization of the Leu10-C_{α} stereocenter, should occur by ring opening, giving an intermediary aldehyde that can racemize at its α position, and subsequent ring closure to the hemiaminal.







Fig. S7 ¹H NMR spectra (watergate, 600 MHz, KH₂PO₄/H₃PO₄ buffer, H₂O/D₂O 5:1, 0.19 mol/l SDS*d*₂₅, pH 6.6) of the equilibrium between TycA-CH(OH)₂ and Ψ [CH(OH)NH]-TycA at different temperatures (290 to 330 K in 10 K steps). (a) Enlargements of NH region with all resolved signals assigned. The set of signals with increases in intensity with higher temperature is marked by grey bars. (b) Complete spectra of TycA-CH(OH)₂/ Ψ [CH(OH)NH]-TycA equilibrium system.





Fig. S8 Up-field region of ¹H NMR spectra (watergate, 600 MHz, H₂O/D₂O 5:1, 0.19 mol/I SDS-d₂₅, pH 6.6) of the equilibrium between TycA-CH(OH)₂ and Ψ [CH(OH)NH]-TycA at different temperatures (290 to 330 K in 10 K steps). Spectra of TycA-CH(OH)₂ and Ψ [CH(OH)NH]-TycA are also shown for comparison. The integrals of the Val8 and Leu10 methyl groups were used to determine the ratio of linear and cyclic species.

Estimation of the macrocyclization entropy

The determination of the macrocyclization entropy based on the evaluation of appropriate ¹H NMR signal integrals has been described in detail for the nostocyclopeptides.^{S6} For this purpose, signals have to be identified which are well-resolved and which do not overlap with other signals over the examined temperature range so that integration is enabled. Furthermore, they must be located sufficiently far from the residual H₂O peak as the applied WATERGATE suppression^{S2} also affects the intensity of nearby signals. The integrals of the methyl signals of Val8 and Leu10 (and a Pro2 γ -H signal of Ψ [CH(OH)NH]-TycA) were used to determine the percentages of TycA-CH(OH)₂ and Ψ [CH(OH)NH]-TycA. Fig. S8 shows the respective region of the NMR spectra at all temperatures investigated, together with the percentages of linear and cyclic species as determined from the signal integrals.

The presence of a second cyclic species which is generated from the hemiaminal main product (see p. S12/S13) has to be considered for evaluation. At higher temperatures where this species is present to a significant extent the system was allowed to come to equilibrium also with respect to this transformation (for example, at 330 K not significant change in the ratio was observed over several hours). The Val8 and Leu10 methyl signals appear at the same chemical shifts for both cyclic species and therefore by integration the whole equilibrium system is considered. As in both possible cases of this transformation (see p. S12/S13) the entropy balance is equal to zero, the obtained macrocyclization balance should, in the case of a sufficient time of equilibration prior to the NMR measurement, count for the transformation between TycA-CH(OH)₂ and the hemiaminal main product Ψ [CH(OH)NH]-TycA.

NOE-based molecular modeling of Ψ[*R*-CH(OH)NH]-TycA

The selected backbone NOE contacts and the NOE intensity derived distances (target dist.) are given in Table S4. The molecular dynamics (MD) simulation resulted in distances listed in the MD column. In the final step, the NOE constraints were removed and the structure was allowed to relax to energy minimization (EM). The resulting distances are shown in the EM column. For the MD and EM experiment, all differences to the target distances are listed. The last column shows all changes in distances from the NOE-derived structure to the energy minimized structure (Diff. EM–MD). Except for the distance restraint between the hemiaminal-OH and Phe3-NH no significant differences were observed before and after energy minimization, which speaks for the plausibility of the structure obtained.

NOE	target dist.	MD	Diff.	EM	Diff.	Diff. EM–MD
Asn5 CO / Val8 NH	2.00	2.10	0.10	2.30	0.30	0.20
Asn5 NH / Val8 C <u>O</u>	2.00	2.10	0.10	2.15	0.15	0.05
Phe3 CO / Leu10 NH	2.00	1.90	- 0.10	1.95	- 0.05	0.05
Phe3 NH / Leu10 CH(OH)NH	1.80	2.30	0.50	3.25	1.45	0.95
D-Phe1 α-H / Leu10 C <u>H</u> (OH)NH	2.50	2.65	0.15	2.85	0.35	0.20
Pro2 α -H / Phe3 NH	3.50	3.55	0.05	3.60	0.10	0.05
D-Phe4 α -H / Asn5 NH	2.50	2.45	- 0.05	2.45	- 0.05	0.00
D-Phe4 α -H / Orn9 NH	2.50	2.30	- 0.20	2.20	- 0.30	- 0.10
Tyr7 NH / Val8 NH	2.50	2.60	0.10	2.70	0.20	0.10
Tyr7 α -H / Val8 NH	3.50	3.50	0.00	3.50	0.00	0.00
Orn9 α -H / Leu10 NH	2.20	2.10	- 0.10	2.20	0.00	0.10

Table S4 NOE backbone restraints and acquired distances for Ψ [*R*-CH(OH)NH]-TycA.

References

- S1 F. Kopp, C. Mahlert, J. Grünewald and M. A. Marahiel, *J. Am. Chem. Soc.*, 2006, **128**, 16478–16479.
- S2 T.-L. Wang and A. J. Shaka, J. Magn. Reson., Ser. A, 1995, 112, 275–279.
- S3 M. Haack, S. Enck, H. Seger, A. Geyer and A. G. Beck-Sickinger, J. Am. Chem. Soc., 2008, 130, 8326–8336.
- S4 S. Laiken, M. Printz and L. C. Craig, J. Biol. Chem., 1969, 244, 4454–4457.
- S5 X. Bu, X. Wu, G. Xie and Z. Guo, Org. Lett., 2002, 4, 2893–2895.
- S6 S. Enck, F. Kopp, M. A. Marahiel and A. Geyer, *ChemBioChem*, 2008, 9, 2597–2601.