

ARTICLE

Received 4 Sep 2013 | Accepted 15 Nov 2013 | Published 12 Dec 2013

DOI: 10.1038/ncomms3945

A very stable complex of a modified marine cyclopeptide with chloroform

Gebhard Haberhauer¹, Áron Pintér¹ & Sascha Woitschetzki¹

Noncovalent interactions play a pivotal role in molecular recognition. These interactions can be subdivided into hydrogen bonds, cation- π interactions, ion pair interactions and London dispersion forces. The latter are considered to be weak molecular interactions and increase with the size of the interacting moieties. Here we show that even the small chloroform molecule forms a very stable complex with a modified marine cyclopeptide. By means of high-level quantum chemical calculations, the size of the dispersive interactions is calculated; the dispersion energy (approximately $-40 \text{ kcal mol}^{-1}$) is approximately as high as if the four outer atoms of the guest form four strong hydrogen bonds with the host. This strong binding of chloroform to a modified marine cyclopeptide allows the speculation that the azole-containing cyclopeptides-haloform interaction may play some biological role in marine organisms such as algae.

¹Institut für Organische Chemie, Universität Duisburg-Essen, Universitätsstrasse 7, D-45117 Essen, Germany. Correspondence and requests for materials should be addressed to G.H. (email: gebhard.haberhauer@uni-due.de).

Molecular recognition is a key process in chemistry and biology. Of high interest are noncovalent host–guest interactions that are mainly dominated by London forces^{1–7}. Here an outstanding challenge is the design of host molecules which are able to bind small nonpolar guests, because these types of guests exhibit only weak dispersion interactions owing to their small number of atoms^{8–10}. A prominent example for a host system, which includes small nonpolar guests, is the cryptophane family^{11,12}. Cryptophane cages consist of two C₃-symmetric cup-shaped cyclotrimeratrylene units (**1**, see Fig. 1), which are connected via three mostly aliphatic linkers. Cryptophane-A (**5**) is able to bind chloroform with an association constant K_i of 230 M^{−1} at 298 K in C₂D₂Cl₄ (refs 13,14). The by far highest affinity ($K_i = 3,000 \text{ M}^{-1}$) has been measured for the Xe@cryptophane-A complex¹⁵. It should be noted that the complex formation constants were measured under conditions excluding hydrophobic effects. If these are not excluded, even higher binding constants are found. For example, a water-soluble triacetate-functionalized cryptophane-A derivative shows the highest affinity for xenon, which was ever measured for a host molecule ($K_i = 33,000 \text{ M}^{-1}$ at 293 K in buffer)¹⁶. An increase of the association constant can also be obtained by adjusting the cavity to the guest¹⁷. For example, cryptophane-E (**6**) has a slightly larger cavity than cryptophane-A and therefore binds chloroform with a higher association constant ($K_i = 470 \text{ M}^{-1}$ at 300 K in C₂D₂Cl₄)^{18,19}.

Recently, we dealt with modified marine cyclopeptides (for example, **2–4**), which exhibit three azole units in their cyclopeptidic scaffold and are thus structurally similar to cyclotrimeratrylene (Fig. 1)^{20,21}. The naturally occurring azole cyclopeptides are mainly biosynthesized by cyanobacteria (blue-green algae); the azoles arise from natural amino-acid residues by enzyme-mediated post-translational modification²². The *in vivo* role of these azole-containing cyclopeptides is still unknown, but because of their size and conformation it has been speculated that they are possibly used for metal complexation or as transporting agents²³. We could show that the spatial structure of the azole-containing macrocycle is strongly dependent on the azole ring²⁴. However, in all cases only one conformation exists in solution; this means that an inversion of the macrocycle does not take place. Furthermore, structural investigations of modified marine cyclopeptides revealed that small halogenated hydrocarbons, such as methylene chloride or chloroform, are enclosed into the solid²⁵. They are mainly found within the concave cavity, which is formed by the three azole rings.

Here we report the properties of the C₃-symmetric imidazole-containing cyclopeptide cavitand **7** (Fig. 1) as host for chloroform and other 1,1,1-trihalomethanes. One cup-shaped side of container **7** is formed by the imidazole-containing cyclopeptide **4**, which is connected via amide bonds to a second cup consisting of a triphenylphosphane oxide unit. NMR investigations reveal that container **7** forms an extremely stable complex with chloroform. High-level quantum chemical calculations show that the dispersive interactions are responsible for the measured complex stability of CHCl₃@**7**.

Results

Determination of the binding constants. The synthesis of **7** was carried out by established protocols²⁶ in a few steps (Supplementary Fig. S1). For excluding hydrophobic effects and for a better comparison with the already well-known cryptophane compounds **5** and **6** (refs 13,18), NMR complexing studies of cavitand **7** with chloroform were performed in C₂D₂Cl₄. This solvent is too large to fit within the cavity and guarantees that the receptor **7** is empty. The addition of chloroform to container **7** at 298 K leads to the formation of the inclusion compound CHCl₃@**7** and a signal of the enclosed chloroform appears at $\delta = 5.19 \text{ p.p.m.}$ in the ¹H NMR spectrum (Fig. 2a,b). This highfield shift ($\Delta\delta = 2.10 \text{ p.p.m.}$) is consistent with an intramolecular inclusion of the guest into the cavity of **7**. The host signals are also shifted by complexation. If an excess of chloroform is used, the signals of the free host disappear completely (Fig. 2b).

A complete removal of chloroform can only be achieved by heating a diluted solution for several hours in C₂D₂Cl₄ at 413 K. As the equilibrium is so strongly shifted to the side of the complex, the thermodynamic parameters were determined by temperature-dependent NMR spectroscopy between 353 and 398 K in C₂D₂Cl₄ (Supplementary Figs S2, S3 and S4). At these temperatures and millimolar concentrations, it is guaranteed that a sufficient amount of uncomplexed container **7** is present in solution. The inclusion/exclusion process is, even at these temperatures, so slow that separate ¹H NMR signals for the complex, the host and chloroform are obtained (Fig. 2c). For example, integral evaluation at 395 K leads to a binding constant K_i of 420 M^{−1}. The plot of $\ln K_i$ against $1/T$ in the area of 353–398 K furthermore allows the determination of the enthalpy (ΔH_i) and the entropy (ΔS_i) of the inclusion (Table 1). As expected, the inclusion of chloroform by container **7** is strongly enthalpy driven, while the entropy disfavors this process. With the help of enthalpy (ΔH_i) and entropy (ΔS_i), it is

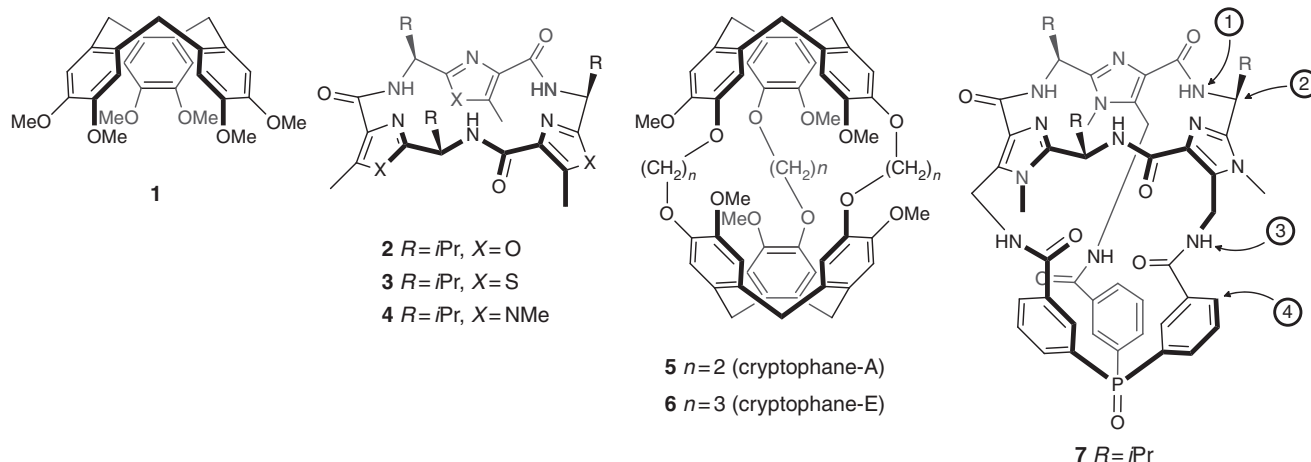


Figure 1 | Host molecules and their building blocks. Structures of cyclotrimeratrylene (**1**), the azole-containing macrocycles **2–4**, cryptophane-A (**5**), cryptophane-E (**6**) and the imidazole-containing container **7**.

also possible to calculate the free enthalpy of inclusion (ΔG_i) and hence the binding constant at 300 K. At this temperature K_i amounts to $140,000 \text{ M}^{-1}$ for $\text{CHCl}_3@7$ and is thus more than two orders of magnitude higher than the binding constants for $\text{CHCl}_3@5$ and $\text{CHCl}_3@6$ (Table 1).

To shed light on the role of the solvent in the complex formation, we tried to determine the binding constant in dichloromethane- d_2 , methanol- d_4 and toluene- d_8 . Unfortunately, temperatures equal to or higher than 353 K are needed to reach the equilibrium within a reasonable timescale and the determination of K_i could not be performed in pure dichloromethane or methanol. Therefore, dichloromethane (100-fold excess over chloroform) was added to the complex $\text{CHCl}_3@7$ in $\text{C}_2\text{D}_2\text{Cl}_4$. In this solvent mixture, K_i was determined to be 900 M^{-1} at 353 K ($K_i = 1,700 \text{ M}^{-1}$ at 353 K in pure $\text{C}_2\text{D}_2\text{Cl}_4$). In other words, dichloromethane reduces the binding constant of chloroform by solvent competition, but the effect is not very pronounced. The binding constant in toluene, which, like $\text{C}_2\text{D}_2\text{Cl}_4$, is too large to fit within the cavity, was measured to be 480 M^{-1} at 359 K (Table 1). The determined enthalpy and the entropy of the inclusion are similar to that found in $\text{C}_2\text{D}_2\text{Cl}_4$ (Supplementary Fig. S5).

To test whether the receptor 7 can also include other 1,1,1-trihalomethanes within its cavity, CHCl_2Br , CHClBr_2 and CHBr_3 were added to a solution of 7 in $\text{C}_2\text{D}_2\text{Cl}_4$. For CHCl_2Br and

CHClBr_2 , no formation of an inclusion compound was observed at room temperature. The inclusion/exclusion process is so slow that temperatures higher than 383 K are needed to reach equilibrium within a reasonable timescale. The binding constants at 395 K amount to 340 M^{-1} for CHCl_2Br and 120 M^{-1} for CHClBr_2 (Table 1). Thus, the binding constants decrease with increasing size of the guest. The obtained enthalpy (ΔH_i) and the entropy (ΔS_i) of the inclusion for CHCl_2Br and CHClBr_2 are in the same range as those found for CHCl_3 (Table 1 and Supplementary Figs S6 and S7). For CHBr_3 , no formation of an inclusion compound could be observed even at 410 K, which is likely due to kinetic reasons (too small entrance into the cavity).

Conformational analysis of the inclusion complexes. The enclosed guest can in principle adopt two orientations, which are as follows: on the one hand, the proton of the trihalomethane can point towards the cup-shaped cyclopeptide; on the other hand, it can point towards the phosphor atom. In the following, the first (second) orientation is labelled with the symbol $\uparrow(\downarrow)$. The orientation can be determined by two-dimensional (2D) nuclear Overhauser enhancement spectroscopy (NOESY) experiments: The ^1H - ^1H NOESY spectrum of $\text{CHCl}_3@7$ shows a strong cross peak between the proton of the enclosed chloroform and the amide protons (H1) of the azole-containing cyclopeptide (Fig. 3; for numbering see Fig. 1), which is an unambiguous proof for the complex $\uparrow\text{CHCl}_3@7$. If the distance between the amide proton H1 and the proton of the adjacent methine group H2 is fixed at 2.94 Å, the value of the $\text{Cl}_3\text{CH} \cdots \text{H1}$ distance amounts to 2.65 Å. This means that the CHCl_3 proton is not only pointing towards the cup-shaped cyclic peptide but is even located deep within the cone of the peptidic scaffold. The ^1H - ^1H NOESY spectra of $\text{CHCl}_2\text{Br}@7$ and $\text{CHClBr}_2@7$ show that here again the proton of the CHX_2Y molecule point towards the cup-shaped cyclic peptide (Supplementary Figs S8 and S9). Also of interest is the fact that the ^1H NMR signals for the complexes $\text{CHCl}_2\text{Br}@7$ and $\text{CHClBr}_2@7$ show C_3 symmetry at 395 K, but not at 298 K (Supplementary Figs S10 and S11). This means that the rotation of the guests is restricted at 298 K because of an energetically favoured position of the guests within the container, which also explains the high loss of entropy by inclusion.

In principle, the triphenylphosphine oxides 7 and $\text{CHCl}_3@7$ can adopt four different conformations ($P1$, $P2$, $M1$ and $M2$): On one side, the three phenyl rings can be present in two opposite helicities (P or M isomer) and on the other side, the amide protons H3 can point towards ($P1$ and $M1$) or away from

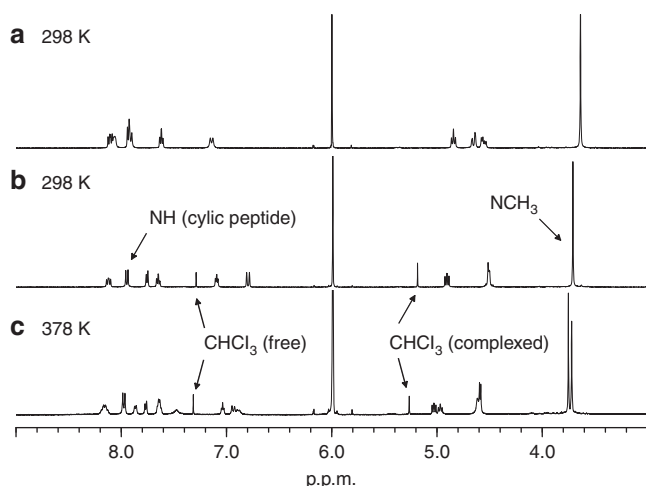


Figure 2 | ^1H NMR spectra of 7 and $\text{CHCl}_3@7$. The ^1H NMR spectra of 7 (a) and $\text{CHCl}_3@7$ (b,c) were measured at 500 MHz in $\text{C}_2\text{D}_2\text{Cl}_4$.

Table 1 | Thermodynamic data determined by ^1H NMR spectroscopy.

Host	Guest	$\Delta\delta$	Solvent	T (K)	K_i (M^{-1})	ΔG_i	ΔH_i	ΔS_i
5 ¹³	CHCl_3	4.33	$\text{C}_2\text{D}_2\text{Cl}_4$	298	230	−3.2	−8.2	−16
6 ¹⁸	CHCl_3	4.44	$\text{C}_2\text{D}_2\text{Cl}_4$	300	470	−3.7	−6.9	−11
	CHCl_2Br	4.42			300	−3.4	−5.2	−6
	CHClBr_2	4.41			130	−2.9	−1.5	+4
7	CHCl_3	2.10	$\text{C}_2\text{D}_2\text{Cl}_4$	300	$140,000 \pm 30,000$	−7.1 [†]	−16 ± 1	−29 ± 2
		2.10		353	$1,700 \pm 90$	−5.2		
		0.89	$\text{C}_6\text{D}_5\text{CD}_3$	359	480 ± 20	−4.4	−16 ± 1	−32 ± 3
		2.10	$\text{C}_2\text{D}_2\text{Cl}_4$	395	420 ± 15	−4.2		
	CHCl_2Br	2.12	$\text{C}_2\text{D}_2\text{Cl}_4$	395	340 ± 15	−3.9	−17 ± 2	−32 ± 3
	CHClBr_2	2.21	$\text{C}_2\text{D}_2\text{Cl}_4$	395	120 ± 5	−3.1	−16 ± 2	−30 ± 3
	CHBr_3		$\text{C}_2\text{D}_2\text{Cl}_4$	395	‡			
					‡			
					‡			

¹³ ^1H NMR highfield shift of the 1,1,1-trihalomethanes signal in the inclusion complex ($\Delta\delta = \delta_i - \delta_o$), binding constant (K_i), free enthalpy (ΔG_i in kcal mol^{-1}), enthalpy (ΔH_i in kcal mol^{-1}) and entropy (ΔS_i in $\text{cal mol}^{-1} \text{K}^{-1}$) for the inclusion reaction. All data were determined by temperature-dependent ^1H NMR measurements in $\text{C}_2\text{D}_2\text{Cl}_4$.

*Calculated based on ΔG_i at 300 K.

†Calculated based on ΔH_i and ΔS_i at 300 K.

‡No complex formation was observed.

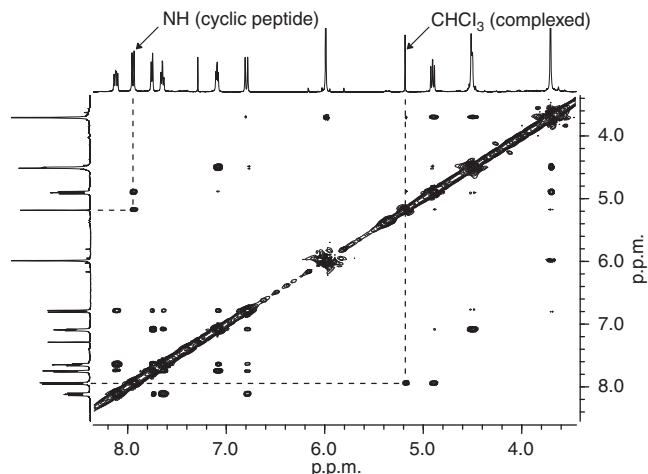


Figure 3 | Part of the ^1H - ^1H NOESY NMR spectrum of $\text{CHCl}_3@7$. The NMR spectrum was measured at 500 MHz in $\text{C}_2\text{D}_2\text{Cl}_4$ (298 K). The cross peaks between the proton of the complexed chloroform and the amide protons of the cup-shaped imidazole scaffold are highlighted.

(*P2* and *M2*) the protons H4 of the phenyl rings (for numbering see Fig. 1, see also Supplementary Fig. S12). To determine the conformations of the free container **7** and the inclusion compound $\text{CHCl}_3@7$, 2D NOESY and circular dichroism measurements were carried out. For both compounds, cross peaks between the amide protons H3 and the protons H4 of the phenyl rings are found in the ^1H - ^1H NOESY NMR spectra (Supplementary Figs S13 and S14). Furthermore, the circular dichroism spectra of both compounds show negative Cotton effects at 265 nm and positive Cotton effects at 250 nm, which arise from *P* helicity of the triphenylphosphine oxide chromophore (Supplementary Fig. S15)²⁶. These data suggest that the inclusion of chloroform does not lead to a conformational change of the container and both molecules adopt *P1* conformation in solution. To support this suggestion, quantum chemical calculations were performed. The structures of all conformers were determined by geometry optimizations at the density functional theory (DFT) level using the MO5-2X functional²⁷. The calculations confirm the assumption that only the *P1* diastereomers are present in solution (Supplementary Figs S16 and S17). A superimposition of the calculated molecular structures of the empty and the chloroform-filled receptor shows that they are very similar (Supplementary Fig. S18).

Calculation of the attractive interaction within $\text{CHCl}_3@7$. To evaluate the contribution of attractive dispersion forces to the highly negative enthalpy of the inclusion, *ab initio* calculations were carried out. The molecular structures were optimized using the DFT potential MO5-2X, which has been shown to cover midrange correlation (Fig. 4)²⁷. For the light elements C, H, N and O the 6-31G* basis set was used, whereas for the heavy elements P, S and the halogens (F, Cl and Br) the correlation-consistent polarized valence triple-zeta (cc-pVTZ) basis set was employed. On the basis of these geometries, single-point calculations with B3LYP^{28–30} and B3LYP-D3 were carried out. While the first method cannot describe the dispersion interaction correctly, especially not at large distances, the latter one includes an additional dispersion correction^{31–33}. Thus, the difference between the two is a hint to the extent of the complex dispersion energies. For the light elements C, H, N and O, the 6-311++G** basis set was used, whereas for the heavy elements P, S and the halogens (F, Cl and Br) the aug-cc-pVTZ basis set was employed.

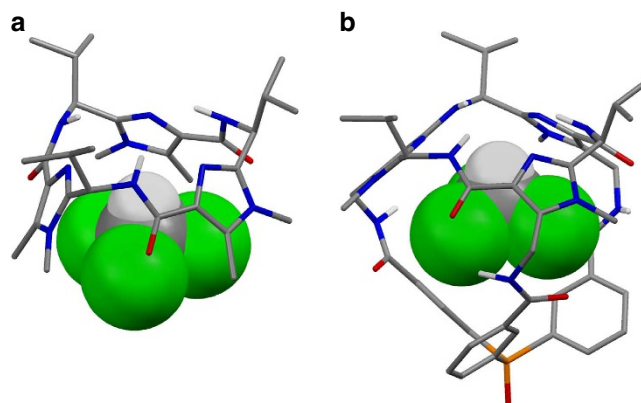


Figure 4 | Molecular structures of $\uparrow\text{CHCl}_3@4$ and $\uparrow\text{CHCl}_3@7$. The geometry of the complexes $\uparrow\text{CHCl}_3@4$ (a) and $\uparrow\text{CHCl}_3@7$ (b) was optimized using MO5-2X/6-31G*,cc-pVTZ. Most hydrogen atoms are omitted for clarity.

The so obtained data are summarized in Table 2 (see also Supplementary Table S1 and Supplementary Data 1).

A glance at the table reveals that inclusion of $\text{C}_2\text{H}_2\text{Cl}_4$ by **7** is energetically disfavoured and thus the use of $\text{C}_2\text{D}_2\text{Cl}_4$ as solvent guarantees the exclusion of the solvent from the cavity. As mentioned above, the enclosed chloroform can adopt two orientations. A comparison of the calculated values shows that all methods favour complex $\uparrow\text{CHCl}_3@7$. The distance between the proton of the enclosed chloroform and the amide protons of the azole-containing cyclopeptide (H1) is calculated to be 2.886 Å in $\uparrow\text{CHCl}_3@7$; in the energetically disfavoured complex $\downarrow\text{CHCl}_3@7$ the distance amounts to 5.712 Å. The first value is consistent with the one found by ^1H - ^1H NOESY NMR measurements (2.65 Å).

The binding energies for the complex $\uparrow\text{CHCl}_3@7$ calculated using MO5-2X and B3LYP-D3 ($-20.5 \text{ kcal mol}^{-1}$ and $-26.5 \text{ kcal mol}^{-1}$) are however significantly higher than the experimentally determined enthalpy of $-16 \text{ kcal mol}^{-1}$. However, it has to be considered that all calculations were carried out for the gas phase, while the experiments were performed in solution ($\text{C}_2\text{D}_2\text{Cl}_4$). Assuming that the interactions of the host and the complex with the solvent are similar, only the interaction between chloroform and the solvent molecules has to be taken into account. To simulate this interaction, we calculated the complex formation of two molecules $\text{C}_2\text{H}_2\text{Cl}_4$ and one molecule chloroform (see Table 2). If this interaction energy is subtracted from the complex formation energy of $\uparrow\text{CHCl}_3@7$, the revised values $-15.3 \text{ kcal mol}^{-1}$ (MO5-2X) and $-20.2 \text{ kcal mol}^{-1}$ (B3LYP-D3) are obtained, which are now quite similar to the experimentally determined enthalpy of $-16 \text{ kcal mol}^{-1}$.

Table 2 shows that the energy values for the complex $\uparrow\text{CHBr}_3@7$ calculated using MO5-2X and B3LYP-D3 (-19.8 and $-29.5 \text{ kcal mol}^{-1}$) are similar to those found for the complex $\uparrow\text{CHCl}_3@7$. This supports the above-made assumption that the complex formation of $\text{CHBr}_3@7$ at 410 K is prevented by kinetic reasons and not by thermodynamics. According to the calculations, the cyclic peptides **2–4** should also form stable complexes with chloroform. Calculated gas-phase interaction energies are approximately from -12 to $-13 \text{ kcal mol}^{-1}$ (MO5-2X) and $-16 \text{ kcal mol}^{-1}$ (B3LYP-D3). These high values indicate that the cup-shaped imidazoles in **7** are crucial for the formation of the stable complex $\uparrow\text{CHCl}_3@7$. Furthermore, for imidazole-containing cyclopeptide **4**, complex energy

Table 2 | Calculated interaction energies.

Host	Guest	Orientation	ΔE_i (MO5-2X)*	ΔE_i (B3LYP)†	ΔE_i (B3LYP-D3)‡
7	CHCl ₃	↑	−20.5	+7.3	−26.5
	CHCl ₃	↓	−16.2	+10.1	−23.0
	CHBr ₃	↑	−19.8	+14.6	−29.5
	C ₂ H ₂ Cl ₄		+9.2	+47.8	+2.2
4	CHCl ₃	↑	−13.3	+4.7	−15.9
	CHCl ₃	↓	−11.9	+4.6	−14.1
	CHF ₃	↑	−12.9	+2.1	−11.2
	CHF ₃	↓	−5.8	+4.3	−5.3
	CHBr ₃	↑	−12.6	+6.3	−17.2
	CHBr ₃	↓	−12.8	+5.5	−16.9
	CHCl ₃	↑	−12.0	+2.0	−15.7
	CHCl ₃	↑	−11.6	+3.0	−15.6
2	CHCl ₃		−5.2	+1.7	−6.3
3	CHCl ₃				
2* C ₂ H ₂ Cl ₄	CHCl ₃				

Calculated gas-phase interaction energies ΔE_i of the host-guest complexes of azole-containing cyclopeptides **2-4** and peptidic container **7** with different trihalomethanes. All values are given in kcal mol^{−1}. In the orientation (↑) the chloroform proton points towards (away from) the amide protons of the cup-shaped azole-containing cyclopeptide.
 MO5-2X/6-31G, cc-pVTZ//MO5-2X/6-31G*, cc-pVTZ.
 †B3LYP/6-311 + + G**, aug-cc-pVTZ//MO5-2X/6-31G*, cc-pVTZ.
 ‡B3LYP-D3/6-311 + + G**, aug-cc-pVTZ//MO5-2X/6-31G*, cc-pVTZ.

calculations with CHF₃ and CHBr₃ were carried out. The result is that the interactions are similarly strong and that the orientation of the haloform molecules does not play a major role with the exception of CHF₃.

A comparison of the binding energies for the complex ↑CHCl₃@**7** calculated using MO5-2X and B3LYP-D3 (−20.5 and −26.5 kcal mol^{−1}) shows that they are significantly higher than those values calculated for the complex formation of chloroform with other arenes. Coupled cluster [CCSD(T)] calculations yield binding energy of −5.46 kcal mol^{−1} for the chloroform-benzene dimer³⁴. A van der Waals density functional study finds a binding energy of −8.22 kcal mol^{−1} for the adsorption of chloroform on graphene³⁵. Attraction in these arene-chloroform complexes mainly derives from the dispersion interaction, whereas the electrostatic contribution is small. A comparison between the interaction energies for ↑CHCl₃@**7** calculated with B3LYP and B3LYP-D3 shows that here again the dispersion energy is the crucial term for complex formation. Thus, energy of +7.3 kcal mol^{−1} is even needed for complex formation of ↑CHCl₃@**7** according to B3LYP calculations. Only the consideration of the dispersion correction D3, which amounts to 33.8 kcal mol^{−1} leads to a stabilization of the complex ↑CHCl₃@**7**. The exact extent of dispersion interactions between the monomers in a complex can be obtained via density fitting DFT-symmetry-adapted perturbation theory (DF-DFT-SAPT) calculations^{36–38}. However, very large basis sets near saturation (cc-pVQZ, and so on) are necessary for accurate evaluation of the dispersion energy³⁹. It is thus not possible to calculate the dispersion energy for the entire system ↑CHCl₃@**7**. Therefore, we chose the following two approaches to make a semiquantitative determination of the dispersion energy: for the first approach, we selected a small subsystem, calculated with a very large basis set and extrapolated to the entire system. For the second approach, we chose a large subsystem, which is very similar to ↑CHCl₃@**7**, calculated with a medium-size basis set and extrapolated to the result of the very large basis set.

As a starting point for the first approach, we chose the MO5-2X-optimized structure of ↑CHCl₃@**7**, removed two-thirds of the C₃-symmetric host as well as the isopropyl group and saturated the unsaturated valencies with hydrogen atoms. The isopropyl group should have no influence on the dispersion energy because it is not a part of the cavity. The obtained model complex CHCl₃@**8** is depicted in Fig. 5 and should exhibit one-third of the dispersion interactions of system ↑CHCl₃@**7**. The DF-DFT-

SAPT-calculated dispersion energies for complex CHCl₃@**8** are listed in Table 3 (see also Supplementary Table S2). In all cases, the functional of Perdew, Burke and Ernzerhof (PBE) was used^{40,41}. As expected, the dispersion energy depends on the used basis set. The largest basis set (aug-cc-pVTZ for C, H, N, O and aug-cc-pVQZ for Cl, P) delivers a value of $E_{\text{Disp}} = -14.1$ kcal mol^{−1}. After extrapolation to the system ↑CHCl₃@**7** the dispersion energy amounts to −42 kcal mol^{−1} (3×-14.1 kcal mol^{−1}).

Starting point for the second approach of dispersion energy determination in system ↑CHCl₃@**7** is also the MO5-2X-optimized structure of ↑CHCl₃@**7**. On the basis of this structure, all isopropyl groups as well as three methyl groups were removed, and unsaturated valencies were saturated with hydrogen atoms. The resulting complex CHCl₃@**10** (Fig. 5) was calculated with DF-DFT-SAPT. The cc-pVDZ basis set was used for the light elements and the aug-cc-pVQZ basis set for the heavy elements (Table 3). Therefore, an energy value of −33.7 kcal mol^{−1} is obtained for the dispersion energy. Here it has to be considered that the used basis set does not include the whole dispersion energy. The results for system CHCl₃@**8** show (Table 3) that only 87% of the dispersion energy calculated with aug-cc-pVTZ, aug-cc-pVQZ are received using the cc-pVDZ, aug-cc-pVQZ basis set. If the dispersion energy of CHCl₃@**10** calculated with cc-pVDZ, aug-cc-pVQZ is extrapolated to the larger basis set, the resulting energy value amounts to −39 kcal mol^{−1}.

Both approaches therefore lead to a dispersion energy of about −40 kcal mol^{−1} in the complex ↑CHCl₃@**7**. The results for model complex CHCl₃@**9** (Fig. 5 and Table 3) furthermore show that the main dispersion energy contribution is caused by the interaction of chloroform and the imidazole-containing macrocycle.

Discussion

The C₃-symmetric cavitand **7** consisting of a cup-shaped imidazole-containing cyclopeptide, which is connected to a triphenylphosphane unit via amide linkers, was successfully synthesized and characterized. Temperature-dependent NMR investigations showed that this container forms an extremely stable complex with chloroform. The formation constant for this complex is even more than two orders of magnitude higher than the one found for CHCl₃@cryptophane systems. A reason for the extremely strong complex CHCl₃@**7** is the negative enthalpy

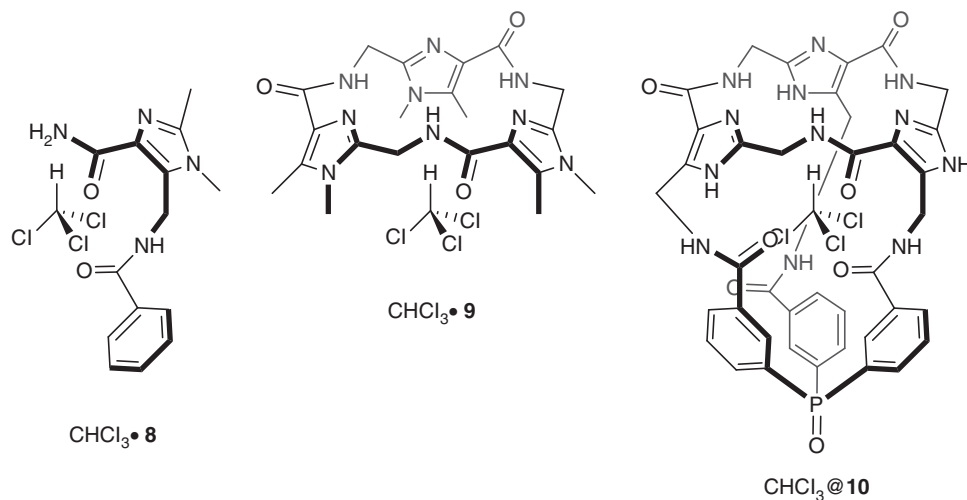


Figure 5 | Model complexes. The model complexes CHCl₃•8, CHCl₃•9 and CHCl₃@10 were used for determination of dispersion energy of complex CHCl₃@7.

Table 3 | Calculated dispersion energies.

Complex	Basis set (C, H, N, O)	Basis set (Cl, P)	E_{disp}
CHCl ₃ •8	cc-pVDZ	aug-cc-pVTZ	−11.5
	cc-pVDZ	aug-cc-pVQZ	−12.2
	cc-pVTZ	aug-cc-pVQZ	−13.3
	aug-cc-pVTZ	aug-cc-pVQZ	−14.1
CHCl ₃ •9	cc-pVDZ	aug-cc-pVQZ	−19.0
CHCl ₃ @10	cc-pVDZ	aug-cc-pVQZ	−33.7

Dispersion energies E_{disp} from DF-DFT-SAPT calculations for the host-guest complexes CHCl₃•8, CHCl₃•9 and CHCl₃@10. All values are given in kcal mol^{−1}.

whose absolute value is twice as high as that for CHCl₃@5 and CHCl₃@6 (Table 1). *Ab initio* calculations (DFT and DF-DFT-SAPT) revealed that a very high dispersion energy (about −40 kcal mol^{−1}) is responsible for the measured complex stability of CHCl₃@7. As chloroform has only four interacting atoms (3 × Cl and 1 × H), a dispersion contribution of about −10 kcal mol^{−1} is found for each outer atom of the guest. It should be mentioned here that strong hydrogen bonds involving neutral species^{42,43} tend to be in the range of 3–15 kcal mol^{−1}. Thus, the attractive dispersive interactions between the guest and the host in ↑CHCl₃@7 equal the energy in a scenario where all four outer atoms of the guest form strong hydrogen bonds to the host.

An explanation for the high dispersion energy is probably the fact that the trigonal bipyramidal cavity of **7** perfectly fits to the shape of the trihalomethanes. The guests are embedded in such a way that a multitude of dispersion interactions between the trihalomethanes and the atoms of the cavity takes place (Supplementary Fig. S19). The surfaces of the guests and the cavity match in size and shape so perfectly that there is not sufficient space for free rotation of the guests. This is confirmed by the restricted rotation of the guests at 298 K. An analogous effect for cryptophanes showing more spherical cavities has not been observed. A further possible reason for the high stability of CHCl₃@7 compared with that of CHCl₃@5 and CHCl₃@6 may be that the cryptophanes do not completely exclude the solvent and thus the container **7** shows improved size discrimination.

The fact that the cup-shaped azole-containing cyclopeptide is mainly responsible for the strong binding of chloroform shows

that these cyclopeptides are new recognition units for 1,1,1-trihaloalkanes and prompt speculation that the azole-containing cyclopeptide–haloform interaction might play some biological role. This speculation is supported by the fact that algae and other marine organisms are the largest source of chloroform, bromoform and other biogenic 1,1,1-trihaloalkanes^{44,45}.

Methods

General. All chemicals were reagent grade and were used as purchased. Reactions were monitored by thin layer chromatography analysis with silica gel 60 F254 thin-layer plates. Flash chromatography was carried out on silica gel 60 (230–400 mesh). ¹H, ¹³C NMR and 2D NMR spectra were measured with Bruker Avance DRX 500 and HD 600 spectrometers. All chemical shifts (δ) are given in p.p.m. The spectra were referenced to deuterated solvents indicated in brackets in the analytical data. HR-MS spectra were recorded with a Bruker BioTOF III Instrument. The cyclic peptide **11** (ref. 26) and the tricarboxylic acid chloride **12** (ref. 46), which are starting materials for the synthesis of **7** (Supplementary Fig. S1), were synthesized as reported.

Synthesis of 7. To a slurry of cyclic peptide **11** (73 mg, 0.10 mmol) in dry dichloromethane (DCM) (30 mL) a 0.50-M stock solution of **12** in DCM (240 μL, 0.12 mmol) was added at 0–5 °C followed by the slow addition of a 1.00-M stock solution of Et₃N in dry DCM (1,000 μL, 1.00 mmol). The cooling bath was then removed, and the mixture was stirred for further 16 h at ambient temperature. Solvent was removed in a rotary evaporator in vacuo, and purification of the product was performed by liquid chromatography on silica gel (CH₂Cl₂/ethyl acetate/methanol 75/25/0 → 75/25/15) to yield 18 mg (18%) of **7** as a colourless solid. ¹H NMR (500 MHz, CDCl₃): δ 8.15 (ddt, $J_1 = 11.4$ Hz, $J_2 = 7.9$ Hz, $J_3 = 1.3$ Hz, 3 H, -ArH), 7.96 (d, $J = 8.8$ Hz, 3 H, NH), 7.70 (dq, $J_1 = 7.9$ Hz, $J_2 = 1.3$ Hz, 3 H, -ArH), 7.61 (td, $J_1 = 7.6$ Hz, $J_2 = 2.2$ Hz, 3 H, -ArH), 7.04 (dd, $J_1 = 7.9$ Hz, $J_2 = 4.4$ Hz, 3 H, NH), 6.81 (dt, $J_1 = 13.2$ Hz, $J_2 = 1.3$ Hz, 3 H, -ArH), 4.94 (dd, $J_1 = 10.4$, $J_2 = 8.2$ Hz, 3 H, CH), 4.55 (dd, $J_1 = 14.8$ Hz, $J_2 = 8.2$ Hz, 3 H, CH₂), 4.49 (dd, $J_1 = 14.8$ Hz, $J_2 = 4.7$ Hz, 3 H, CH₂), 3.75 (s, 9 H, NCH₃), 2.15–2.08 (m, 3 H,

CH), 1.13 (d, $J = 6.6$ Hz, 9H, CH₃), 0.96 (d, $J = 6.6$ Hz, 9H, CH₃); ¹³C NMR (125 MHz, CDCl₃): δ 169.16, 163.29, 148.37, 135.72 ($J_{\text{P,C}} = 12.7$ Hz), 134.89 ($J_{\text{P,C}} = 8.0$ Hz), 133.30 ($J_{\text{P,C}} = 12.2$ Hz), 133.29 ($J_{\text{P,C}} = 105.9$ Hz), 131.05, 130.05 ($J_{\text{P,C}} = 2.1$ Hz), 129.18 ($J_{\text{P,C}} = 11.2$ Hz), 49.7, 34.4, 33.2, 31.0, 19.3, 18.8; ³¹P NMR (202.3 MHz, CDCl₃): δ 33.2 p.p.m.; high-resolution mass spectrometry (electrospray ionization-positive) (m/z): [M + H]⁺ calcd. for C₅₁H₅₈N₁₂O₇P, 981.4289; found, 981.4266; [M + Na]⁺ calcd. for C₅₁H₅₇N₁₂NaO₇P, 1003.4109; found, 1003.4093; [M + CHCl₃ + H]⁺ calcd. for C₅₂H₅₉Cl₃N₁₂O₇P, 1099.3433; found, 1099.3504; [M + CHCl₃ + Na]⁺ calcd. for C₅₂H₅₈Cl₃N₁₂NaO₇P, 1121.3252; found, 1121.3320.

Determination of the thermodynamic parameters. To determine the thermodynamic parameters of CHCl₃@7 in C₂D₂Cl₄, cavitand 7 was first dissolved in chloroform and the solvent was then evaporated. The complex CHCl₃@7 was dissolved in tetrachloroethane-*d*₂ in a concentration of about 1 mM. The NMR samples were heated up to a defined temperature and allowed to equilibrate. According to the ¹H NMR spectra, equilibrium was reached after 2–4 h. The ¹H NMR spectra for determination of the thermodynamic parameters were measured after 6 h and after 1 day. See Supplementary Figs S2, S3 and S4. To determine the thermodynamic parameters of CHCl₃@7 in toluene-*d*₈, cavitand 7 was first dissolved in chloroform and the solvent was then evaporated. The complex CHCl₃@7 was dissolved in toluene-*d*₈ in a concentration of about 1 mM. The NMR samples were heated up to a defined temperature and allowed to equilibrate over night. See Supplementary Fig. S5. To determine the thermodynamic parameters of CHCl₂Br@7 (CHClBr₂@7), cavitand 7 and CHCl₂Br (CHClBr₂) were dissolved in tetrachloroethane-*d*₂ in a concentration of about 1 mM. The NMR samples were heated up to a defined temperature and allowed to equilibrate. All measurements were repeated at least twice. See Supplementary Figs S6 and S7.

Computational study. All calculations were performed by using the program package Gaussian 09 (ref. 47) and MOLPRO⁴⁸. Geometry optimizations were performed using the DFT potential MO5-2X (ref. 27). For the light elements C, H, N and O, the 6-31 G* basis set was used, whereas for the heavy elements P, S and the halogens (F, Cl and Br) the cc-pVTZ basis set was employed. Compounds 2–4 and 7 and their corresponding CHX₃ complexes were completely geometrically optimized within the C₃ point group. The inclusion compound C₂H₂Cl₄@7 was optimized without any symmetry restriction. On the basis of these geometries, single-point calculations with B3LYP (refs 28–30) and B3LYP-D3 were carried out. For the light elements C, H, N and O, the 6-311++G** basis set was used, whereas for the heavy elements P, S and the halogens (F, Cl and Br) the aug-cc-pVTZ basis set was employed. See Supplementary Table S1 and Supplementary Data 1.

Dispersion energies from DF-DFT-SAPT calculations. To assess the contribution of the dispersion energy E_{Disp} in the model complexes CHCl₃@8, CHCl₃@9 and CHCl₃@10, we adopted the terminology derived from the SAPT treatment. In this approach, the total interaction energy E^{SAPT} between closed-shell molecules can be obtained as a sum of individual first and second-order interaction terms, namely, electrostatic $E^{(1)}_{\text{elst}}$, induction $E^{(2)}_{\text{ind}}$ and dispersion $E^{(2)}_{\text{disp}}$ interaction. Each of these terms is accompanied by its exchange counterpart, abbreviated as $E^{(1)}_{\text{exch}}$, $E^{(2)}_{\text{exch-ind}}$ and $E^{(2)}_{\text{exch-disp}}$, respectively. The latter terms arise due to electron exchange between the monomers when the molecules are close to each other. Effects of higher than second order in the intermolecular potential can be estimated from the difference δ (Hartree-Fock; HF) of counterpoise-corrected supermolecular HF calculations and electrostatic and induction energies along with their exchange counterparts obtained from static and response HF-level properties. In sum, using the SAPT treatment the total interaction energy E^{SAPT} between closed-shell molecules can be written as:

$$E^{\text{SAPT}} = E_{\text{elst}} + E_{\text{exch}} + E_{\text{ind}} + E_{\text{Disp}} \quad (1)$$

with

$$E_{\text{ind}} = E^{(2)}_{\text{ind}} + E^{(2)}_{\text{exch-ind}} + \delta(\text{HF}) \quad (2)$$

$$E_{\text{Disp}} = E^{(2)}_{\text{disp}} + E^{(2)}_{\text{exch-disp}} \quad (3)$$

Thus, the dispersion energy, E_{Disp} , can be calculated as the sum of $E^{(2)}_{\text{disp}}$ and $E^{(2)}_{\text{exch-disp}}$. See Supplementary Table S2.

References

1. Steed, J. W. & Atwood, J. L. *Supramolecular Chemistry* (J. Wiley & Sons, 2000).
2. Atwood, J. L. & Steed, J. W. (eds). *Encyclopedia of Supramolecular Chemistry* (Taylor & Francis, 2004).
3. Schneider, H.-J. Binding mechanisms in supramolecular complexes. *Angew. Chem. Int. Ed.* **48**, 3924–3977 (2009).
4. Lehn, J.-M. *Supramolecular Chemistry: Concepts and Perspectives* (J. Wiley & Sons, 1995).
5. Sauvage, J. P. (ed). *Transition Metals in Supramolecular Chemistry. Perspectives in Supramolecular Chemistry* (J. Wiley & Sons, 1999).
6. Vögtle, F. & Alfter, F. *Supramolecular Chemistry: An Introduction* (J. Wiley & Sons, 1993).
7. Salonen, L. M., Ellermann, M. & Diederich, F. Aromatic rings in chemical and biological recognition: energetics and structures. *Angew. Chem. Int. Ed.* **50**, 4808–4842 (2011).
8. Jasat, A. & Sherman, J. C. Carceplexes and hemicarceplexes. *Chem. Rev.* **99**, 931–967 (1999).
9. Cram, D. J. & Cram, J. M. *Container Molecules and Their Guests* (Royal Society of Chemistry, 1994).
10. Conn, M. M. & Rebek Jr., J. Self-assembling capsules. *Chem. Rev.* **97**, 1647–1668 (1997).
11. Gabard, J. & Collet, A. Synthesis of a (*D*₃)-bis(cyclotrimeratrylenyl) macrocage by stereospecific replication of a (C₃)-subunit. *J. Chem. Soc. Chem. Commun.* 1137–1139 (1981).
12. Brotin, T. & Dutasta, J.-P. Cryptophanes and their complexess - present and future. *Chem. Rev.* **109**, 88–130 (2009).
13. Garell, L., Dutasta, J.-P. & Collet, A. Complexation of methane and chlorofluorocarbons by cryptophane-a in organic solution. *Angew. Chem. Int. Ed. Engl.* **32**, 1169–1171 (1993).
14. Marjanska, M., Goodson, B. M., Castiglione, F. & Pines, A. Inclusion complexes oriented in thermotropic liquid-crystalline solvents studied with carbon-13 NMR. *J. Phys. Chem. B* **107**, 12558–12561 (2003).
15. Bartik, K., Luhmer, M., Dutasta, J. P., Collet, A. & Reisse, J. Xe-129 and H-1 NMR study of the reversible trapping of xenon by cryptophane-A in organic solution. *J. Am. Chem. Soc.* **120**, 784–791 (1998).
16. Hill, P. A., Wei, Q., Troxler, T. & Dmochowski, I. J. Substituent effects on xenon binding affinity and solution behavior of water-soluble cryptophanes. *J. Am. Chem. Soc.* **131**, 3069–3077 (2009).
17. Taratula, O., Hill, P. A., Khan, N. S., Carroll, P. J. & Dmochowski, I. J. Crystallographic observation of ‘induced fit’ in a cryptophane host-guest model system. *Nat. Comm.* **1**, 148 (2010).
18. Canceill, J., Lacombe, L. & Collet, A. New cryptophane forming unusually stable inclusion complexes with neutral guests in a lipophilic solvent. *J. Am. Chem. Soc.* **108**, 4230–4232 (1986).
19. Petrov, O., Tosner, Z., Csöreg, I., Kowalewski, J. & Sandström, D. Dynamics of chloromethanes in cryptophane-E inclusion complexes: a 2 H solid-state NMR and X-ray diffraction study. *J. Phys. Chem. A* **109**, 4442–4451 (2005).
20. Haberhauer, G. & Rominger, F. Syntheses and structures of imidazole analogues of *Lissoclinum* cyclopeptides. *Eur. J. Org. Chem.* 3209–3218 (2003).
21. Haberhauer, G. & Pintér, A. Thieme Chemistry Journal Awardees – where are they now? Macrocyclic peptide chemistry inspired by nature – from chiral artificial receptors toward molecular devices. *Synlett* 3082–3098 (2009).
22. Roy, R. S., Gehring, A. M., Milne, J. C., Belshaw, P. J. & Walsh, C. T. Thiazole and oxazole peptides: biosynthesis and molecular machinery. *Nat. Prod. Rep.* **16**, 249–263 (1999).
23. Michael, J. P. & Pattenden, G. Marine metabolites and metal ion chelation: the facts and the fantasies. *Angew. Chem. Int. Ed. Engl.* **32**, 1–23 (1993).
24. Haberhauer, G., Drosdow, E., Oeser, T. & Rominger, F. Structural investigation of westiellamide analogues. *Tetrahedron* **64**, 1853–1859 (2008).
25. Haberhauer, G., Oeser, T. & Rominger, F. Molecular scaffold for the construction of three-armed and cage-like receptors. *Chem. Eur. J.* **11**, 6718–6726 (2005).
26. Pintér, Á., Haberhauer, G., Hyla-Kryspin, I. & Grimme, S. Configurationally stable propeller-like triphenylphosphine and triphenylphosphine oxide. *Chem. Commun.* 3711–3713 (2007).
27. Zhao, Y., Schultz, N. E. & Truhlar, D. G. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2**, 364–382 (2006).
28. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
29. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
30. Miehlich, B., Savin, A., Stoll, H. & Preuss, H. Results obtained with the correlation-energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **157**, 200–206 (1989).
31. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104–154119 (2010).
32. Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **18**, 9955–9964 (2012).
33. Risthaus, T. & Grimme, S. Benchmarking of London dispersion-accounting density functional theory methods on very large molecular complexes. *J. Chem. Theory Comput.* **9**, 1580–1591 (2013).
34. Fujii, A. et al. Experimental and theoretical determination of the accurate interaction energies in benzene-halomethane: the unique nature of the

- activated CH/ π interaction of haloalkanes. *Phys. Chem. Chem. Phys.* **10**, 2836–2843 (2008).
35. Åkesson, J., Sundborg, O., Wahlström, O. & Schröder, E. A van der Waals density functional study of chloroform and other trihalomethanes on graphene. *J. Chem. Phys.* **137**, 174702 (2012).
36. Jeziorski, B., Moszynski, R. & Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **94**, 1887–1930 (1994).
37. Szalewicz, K., Patkowski, K. & Jeziorski, B. Intermolecular interactions via perturbation theory: from diatoms to biomolecules. *Struct. Bond.* **116**, 43–117 (2005).
38. Heßelmann, A., Jansen, G. & Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: a new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **122**, 014103–014117 (2005).
39. Tsuzuki, S., Uchimaru, T., Mikami, M. & Tanabe, K. New medium-size basis sets to evaluate the dispersion interaction of hydrocarbon molecules. *J. Phys. Chem. A* **102**, 2091–2094 (1998).
40. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
41. Perdew, J. P., Burke, K. & Ernzerhof, M. Errata: generalized gradient approximation made simple. *Phys. Rev. Lett.* **78**, 1396 (1997).
42. Emsley, J. Very strong hydrogen bonding. *Chem. Soc. Rev.* **9**, 91–124 (1980).
43. Aakeröy, C. B. *Encyclopedia of Supramolecular Chemistry* (eds Atwood, J. L. & Steed, J. W.) (Taylor & Francis, 2004).
44. Gribble, G. W. The diversity of naturally produced organohalogenes. *Chemosphere* **52**, 289–297 (2003).
45. Gribble, G. W. The diversity of naturally occurring organobromine compounds. *Chem. Soc. Rev.* **28**, 335–346 (1999).
46. Pintér, Á. & Haberhauer, G. Synthesis of chiral threefold and sixfold functionalized macrocyclic imidazole peptides. *Tetrahedron* **65**, 2217–2225 (2009).
47. Pople, J. A. *et al.* Gaussian 09, revision A.02 (Gaussian, Inc., 2009).
48. Werner, H.-J. *et al.* MOLPRO, version 2012.1, <http://www.molpro.net/> (2012).

Acknowledgements

This work was generously supported by the Deutsche Forschungsgemeinschaft (DFG). A special thank goes to Professor Dr. Rolf Gleiter (Heidelberg) for numerous discussions and helpful support. Many thanks also to Patrizia Chamier-Cieminski, Helma Kallweit, Petra Schneider, Dr. Silvia Ernst, Dr. Christoph Wölper and Professor Dr. Georg Jansen (Essen).

Author contributions

G.H. wrote the manuscript and performed the calculations. The synthesis of **7** was conceived and realized by Á.P. The thermodynamic parameters of the complexes were determined by G.H. and S.W.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Haberhauer, G. *et al.* A very stable complex of a modified marine cyclopeptide with chloroform. *Nat. Commun.* **4**:2945 doi: 10.1038/ncomms3945 (2013).