Effects of sulfoxide and sulfone sidechain–backbone hydrogen bonding on local conformations in peptide models

Dayi Liu,[a] Sylvie Robin,[a,b] Eric Gloaguen,[c] Valérie Brenner,[d] Michel Mons*,[e] and David J. Aitken*,[a]

SUPPORTING INFORMATION

Table of Contents

[a] *Université Paris-Saclay, CNRS, ICMMO, 91400 Orsay, France.*
[b] *Université Paris Cité, Faculté de Pharmacie, 75006 Paris, France.*
[c] *Université Paris-Saclay, CNRS, ISMO, 91400 Orsay, France.*
[d] *Université Paris-Saclay, CEA, DRF, 91191 Gif-sur-Yvette, France.*
[e] *Université Paris-Saclay, CEA, LIDYL, 91191 Gif-sur-Yvette, France.*

1. Synthesis

1.1. Materials and instrumentation

Melting points were measured in open capillary tubes on a Büchi B-540 apparatus and were uncorrected. ^{1}H and ^{13}C NMR spectra were recorded at 300 K on a Bruker spectrometer operating at 400 MHz (for ^{1}H) or at 100 MHz (for ^{13}C). For ^{1}H NMR spectra, chemical shifts (δ) are reported in parts per million (ppm) with reference to residual protonated solvent (δ = 7.26 ppm for $CHCl_3$). ^{1}H signals are designated as s (singlet), d (doublet), m (multiplet), or bs (broad singlet); coupling constants (*J)* are reported in hertz (Hz). For ^{13}C NMR spectra, chemical shifts (δ) are reported in parts per million (ppm) with reference to the deuterated solvent (δ = 77.16 ppm for $CDCl_3$). Routine infrared (IR) spectra were recorded for neat solid samples on an FT-IR Perkin Elmer Spectrum Two using an ATR diamond accessory; maximum absorbances (ν) are given in cm^{-1}. High-resolution mass spectra (HRMS [ESI(+)]) were recorded on a Bruker MicroTOF-Q spectrometer equipped with an electrospray ionization source in positive mode.

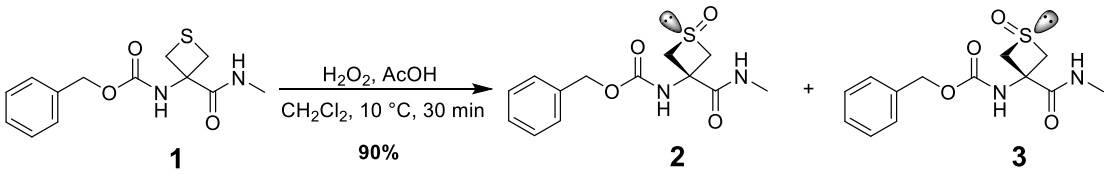
Flash chromatography was performed on columns of silica gel (40-63 μm) purchased from Macherey-Nagel. TLC plates were visualized by fluorescence at 254 nm then revealed using a ninhydrin solution (14 mM in EtOH) or a $KMnO_4$ solution (7.5% in water). Retention factors (R_f) are given for such TLC analysis.

m-Chloroperbenzoic acid (*m*CPBA, 70-75%), H_2O_2 (35 wt.% in water) and potassium carbonate were purchased from Acros; glacial acetic acid was purchased from Carlo Erba; sodium sulphate was obtained from VWR chemicals. All reagents were used as supplied. CH_2Cl_2 was dried by passage through a column of alumina before use.

Compound **1** (Cbz-Attc-NHMe) was prepared following the literature procedure.[1] The spectroscopic data presented in Section S3 are those obtained for samples of **1** made by us for the present study.

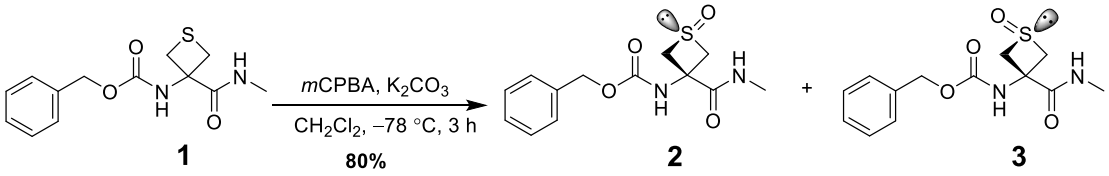
1.2. Synthesis of Cbz-Attc(*trans*-O)-NHMe **2** and Cbz-Attc(*cis*-O)-NHMe **3**

Protocol **A**

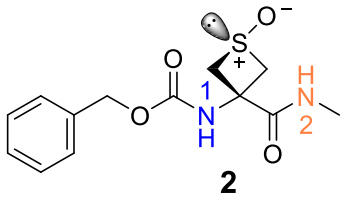


To a cold (10 °C) solution of Cbz-Attc-NHMe **1** (28 mg, 0.1 mmol, 1 eq.) in CH_2Cl_2 (1 mL) in an argon-flushed flask, was added dropwise a mixture of 35% H_2O_2 (40 µL) in glacial acetic acid (40 µL). Then the solution was stirred for 30 min at 10 °C. The reaction mixture was then made alkaline (pH > 10) by addition of a 1M NaOH aqueous solution and extracted with CH_2Cl_2 (5 × 5 mL). The combined organic layers were dried over sodium sulfate, filtered and concentrated under reduced pressure to give the crude product as a white solid. Separation was performed by flash chromatography (eluent EtOAc:CH_2Cl_2:MeOH, gradient from 30:10:0 to 10:10:1) to give Cbz-Attc(*trans*-O)-NHMe **2** (9 mg, 30%) and Cbz-Attc(*cis*-O)-NHMe **3** (17 mg, 60%) as white solids.

Protocol **B**



To a cold (–78 °C) solution of Cbz-Attc-NHMe **1** (28 mg, 0.1 mmol, 1 eq.) in CH_2Cl_2 (4 mL) in argon-flushed flask, potassium carbonate (27.8 mg, 0.2 mmol, 2 eq.) and *m*CPBA (21 mg, 0.1 mmol, 1 eq.) were added successively. The mixture was stirred for three hours at –78 °C and then allowed warm to room temperature. The suspension was filtered through a 0.45 µm PTFE membrane and the filtrate was evaporated under reduced pressure. Separation was performed as above (Protocol A) to give Cbz-Attc(*trans*-O)-NHMe **2** (16 mg, 53%) and Cbz-Attc(*cis*-O)-NHMe **3** (8 mg, 27%) as white solids.



2

R_f = 0.29 (EtOAc:CH_2Cl_2:MeOH = 30:10:1)

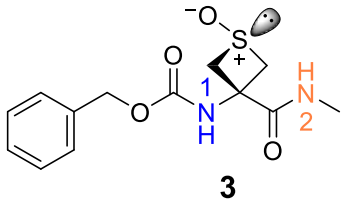
Mp = 190-192 °C

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.78 (bs, 1H, NH^2), 7.43-7.27 (m, 5H, CH^{Ar}), 6.69 (bs, 1H, NH^1), 5.12 (s, 2H, CH_2^{Cbz}), 4.53 (d, J = 14.4 Hz, 2H, $C^{\beta}H^a$), 3.31 (d, J = 14.4 Hz, 2H, $C^{\beta}H^b$), 2.93 (d, J = 4.4 Hz, 3H, N^2CH_3).

^{13}C NMR (100 MHz, $CDCl_3$) δ 170.13 (CO^{amide}), 154.87 (CO^{Cbz}), 136.02 (C^{Ar}), 128.80, 128.54, 128.19 (CH^{Ar}), 67.14 (CH_2^{Cbz}), 59.50 (C^{α}), 56.23 ($C^{\beta}H_2$), 27.45 (N^2CH_3).

IR (neat) ν 3323, 3217, 3033, 1719, 1651, 1541 cm^{-1}.

HRMS [ESI(+)] *m/z* $[M+Na]^+$ calculated for $[C_{13}H_{16}N_2NaO_4S]^+$: 319.0723, found: 319.0716.



3

R_f = 0.24 (EtOAc:CH_2Cl_2:MeOH = 30:10:1)

Mp = 193-195 °C

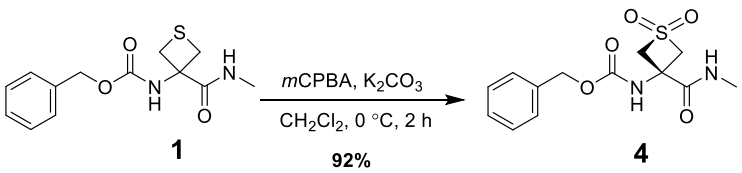
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.46-7.29 (m, 5H, CH^{Ar}), 6.42 (bs, 1H, NH^2), 5.84 (bs, 1H, NH^1), 5.13 (s, 2H, CH_2^{Cbz}), 4.20 (d, J = 12.0 Hz, 2H, $C^{\beta}H^b$), 3.22 (d, J = 12.0 Hz, 2H, $C^{\beta}H^a$), 2.80 (d, J = 4.8 Hz, 3H, N^2CH_3).

^{13}C NMR (100 MHz, $CDCl_3$) δ 171.96 (CO^{amide}), 154.52 (CO^{Cbz}), 135.43 (C^{Ar}), 128.92, 128.92, 128.53 (CH^{Ar}), 68.05 (CH_2^{Cbz}), 61.46 ($C^{\beta}H_2$), 53.13 (C^{α}), 26.90 (N^2CH_3).

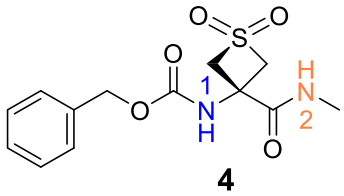
IR (neat) ν 3319, 3227, 3029, 1716, 1637, 1531 cm^{-1}.

HRMS [ESI(+)] *m/z* $[M+Na]^+$ calculated for $[C_{13}H_{16}N_2NaO_4S]^+$: 319.0723, found: 319.0716.

1.3. Synthesis of Cbz-Attc(O,O)-NHMe **4**



To a solution of Cbz-Attc-NHMe **1** (28 mg, 0.1 mmol, 1 eq.) in CH_2Cl_2 (4 mL) in an argon-flushed flask cooled to 0 °C, potassium carbonate (27.8 mg, 0.2 mmol, 2 eq.) and *m*CPBA (21 mg, 0.2 mmol, 2 eq.) were then added successively. The mixture was stirred for two hours at 0 °C and then allowed to warm to room temperature. The suspension was filtered through a sintered glass funnel and the filtrate was evaporated under reduced pressure. Purification was performed by flash chromatography on silica (EtOAc:CH_2Cl_2 = 3:1) to give Cbz-Attc(O,O)-NHMe **4** as a white solid (29 mg, 92%).



***R**f* = 0.78 (EtOAc:CH_2Cl_2: MeOH = 30:10:1)

Mp = 177-179 °C

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58-7.27 (m, 5H, CH^{Ar}), 6.80 (bs, 1H, NH^2), 6.07 (bs, 1H, NH^1), 5.16 (s, 2H, CH_2^{Cbz}), 4.71 (d, *J* = 12.4 Hz, 2H, $C^{\beta}H^b$), 4.33 (d, *J* = 12.4 Hz, 2H, $C^{\beta}H^a$), 2.85 (d, *J* = 4.8 Hz, 3H, N^2CH_3).

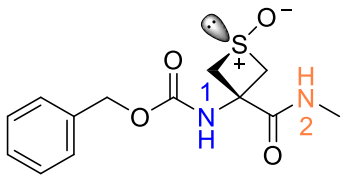
^{13}C NMR (100 MHz, $CDCl_3$) δ 169.11 (CO^{amide}), 155.45 (CO^{Cbz}), 135.38 (C^{Ar}), 128.91, 128.75, 128.50 (CH^{Ar}), 72.90 ($C^{\beta}H_2$), 68.23 (CH_2^{Cbz}), 47.54 (C^{α}), 27.34 (N^2CH_3).

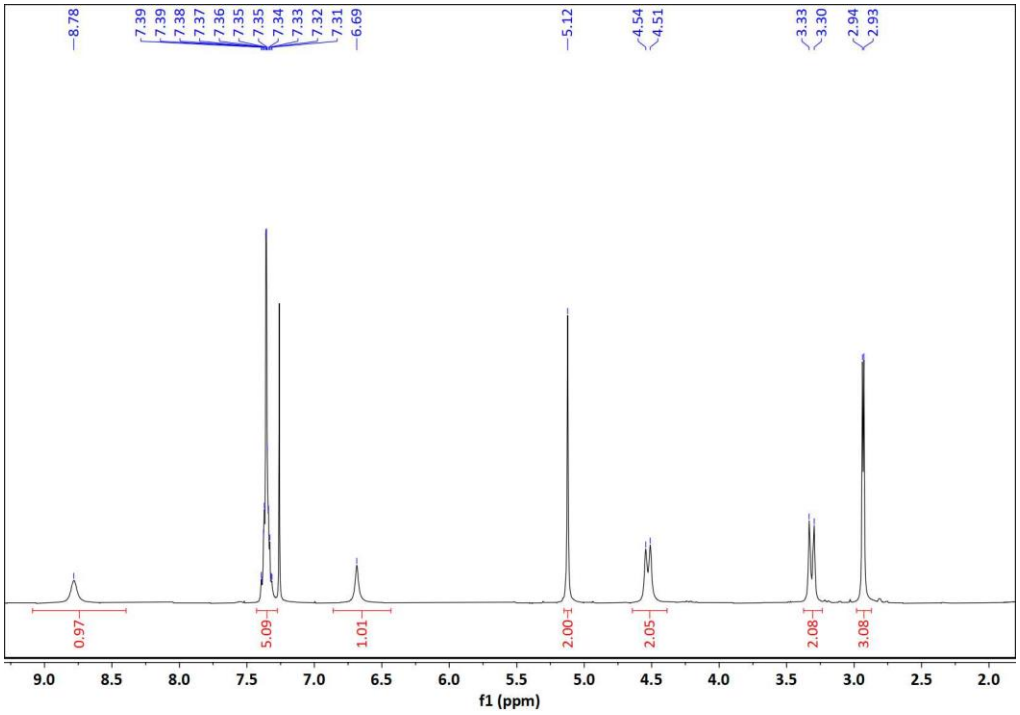
IR (neat) *v* 3383, 3231, 3038, 1725, 1662, 1538 cm^{-1}.

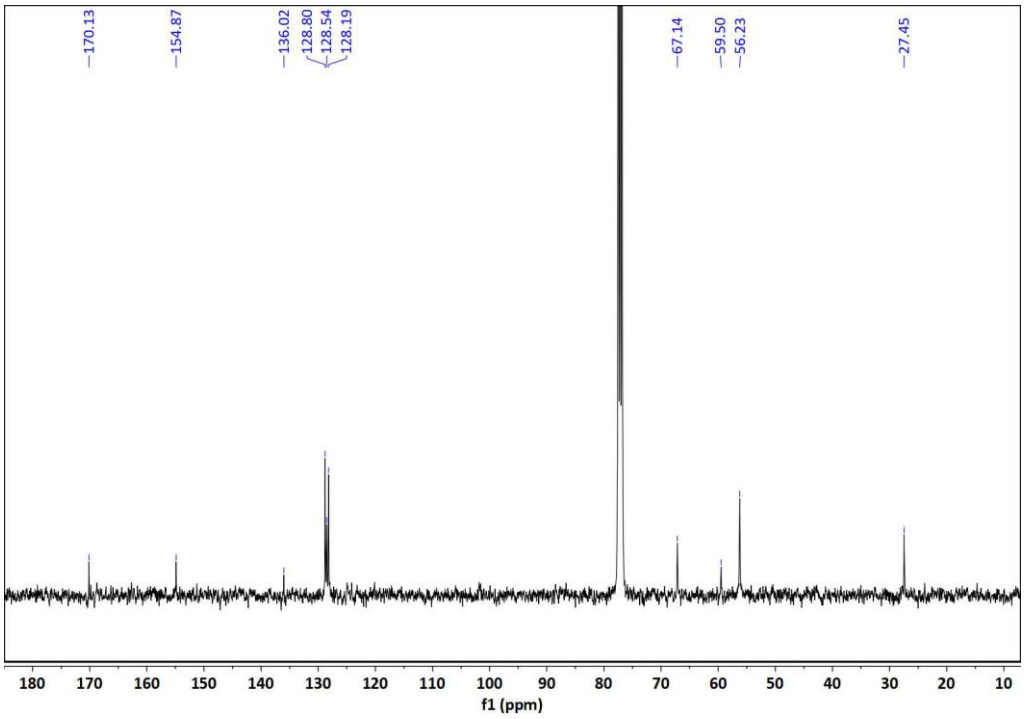
HRMS [ESI(+)] *m/z* $[M+Na]^+$ calculated for $[C_{13}H_{16}N_2NaO_5S]^+$: 335.0672, found: 335.0671.

1.4 Copies of ^{1}H and ^{13}C NMR spectra of compounds **2**-**4**

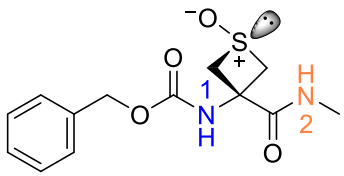
Cbz-Attc(*trans*-O)-NHMe **2**

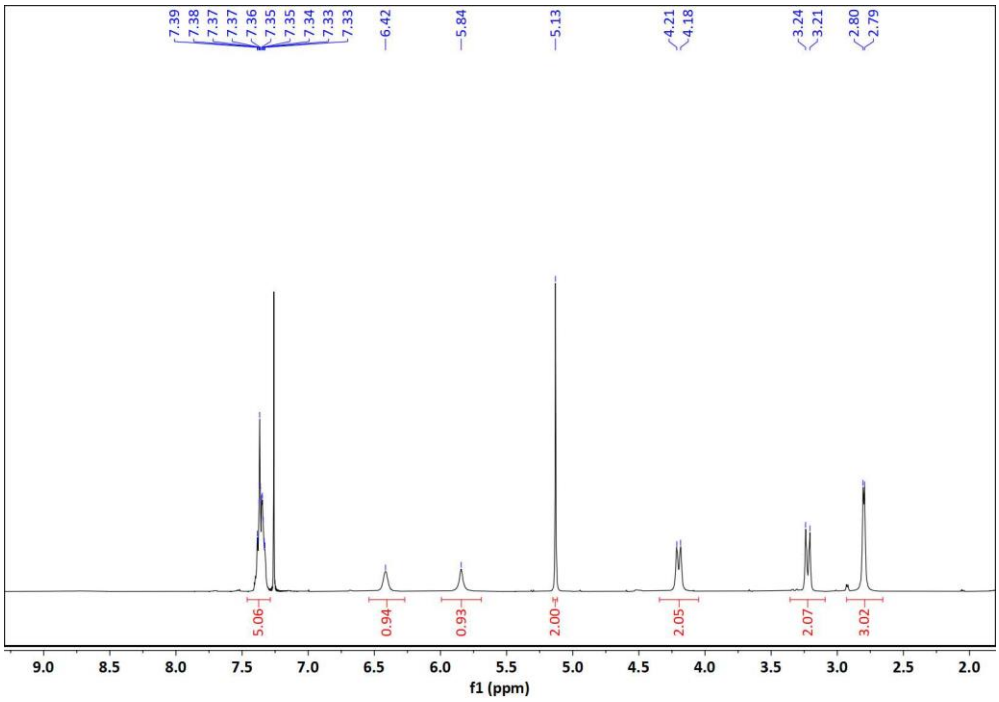


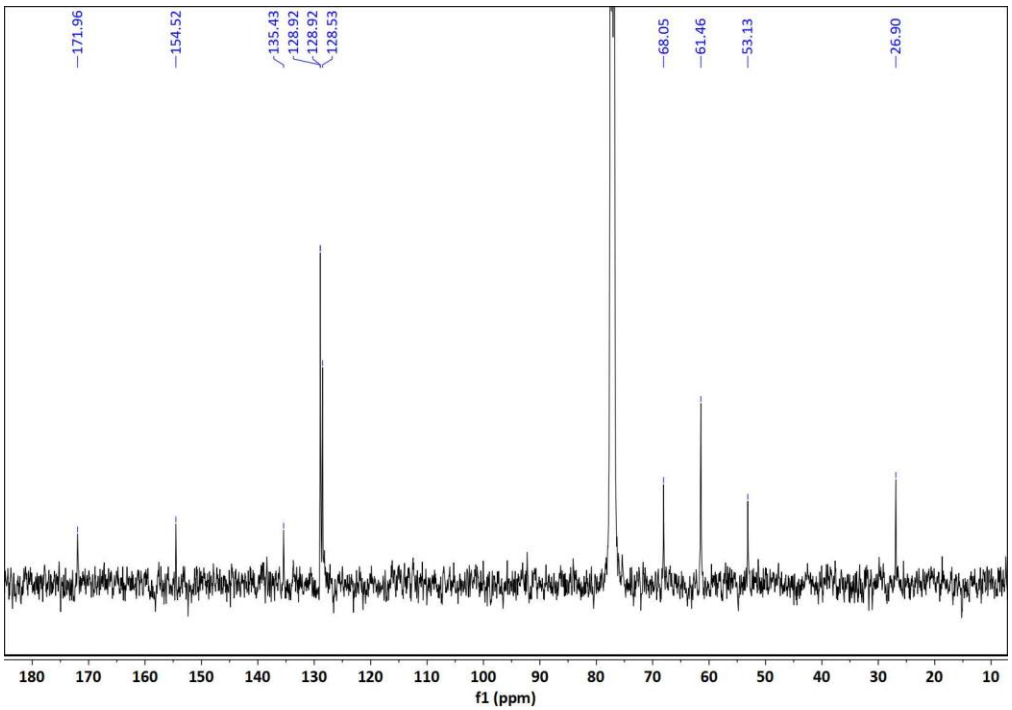




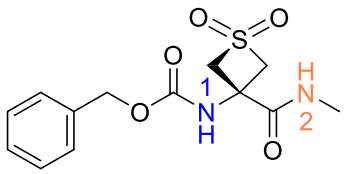
Cbz-Attc(*cis*-O)-NHMe **3**

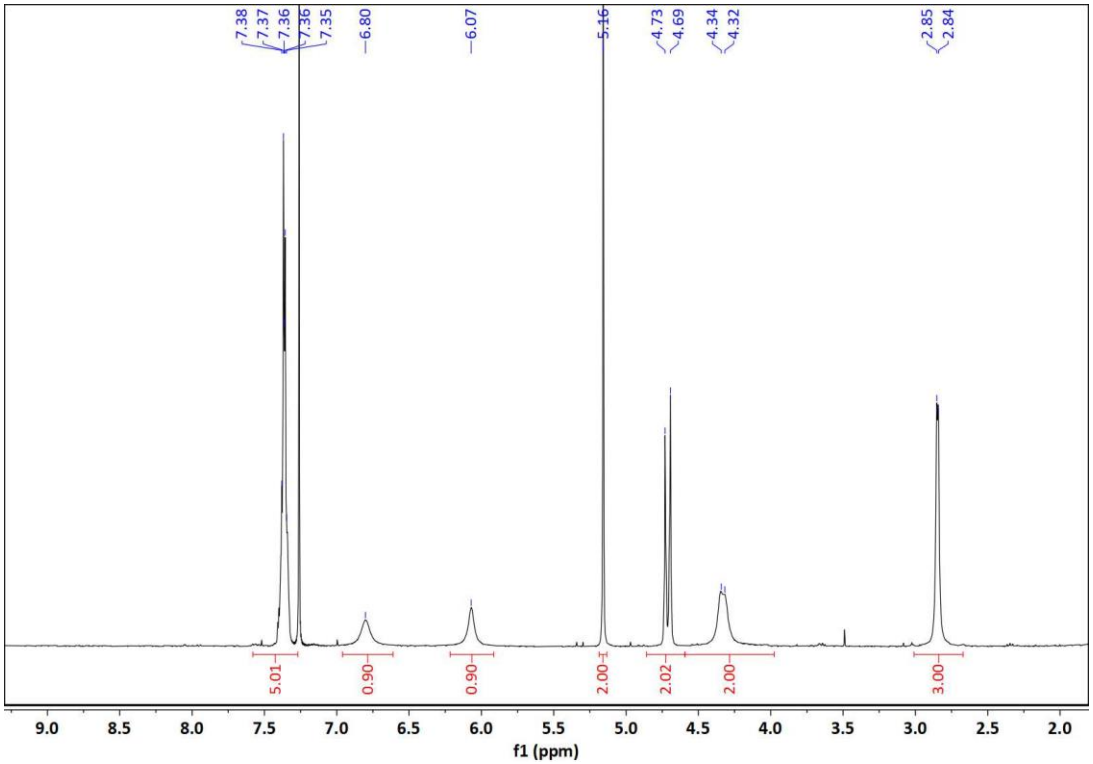


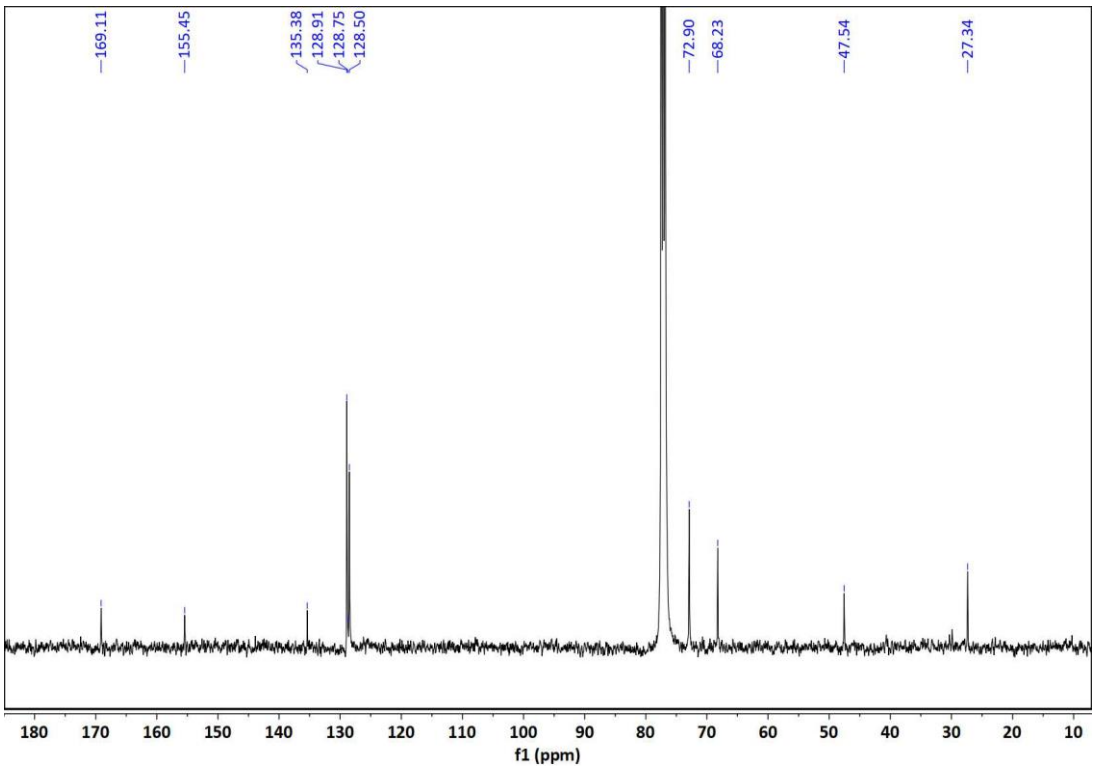




Cbz-Attc(O,O)-NHMe **4**







2. Theoretical chemistry

2.1 Methodology

The theoretical toolbox used was already described in detail in a previous paper.[2] Conformational landscapes of compounds **2**-**4** were constructed in the present work, both in the gas phase and in solution, starting from structures found for compound **1**. These structures were then optimized at a quantum chemistry level (DFT-D), previously validated on compound **1**, for which the theoretical predictions (energetics and scaled vibrational spectra) provided a satisfactory correlation with experiment in both gas phase and solution. For the sake of simplicity, only *gauche* conformations of the Cbz cap were considered since their *trans* counterparts are usually slightly less stable and their orientation does not favor stabilizing interactions with side-chain motifs.

The structures were first optimized in the gas phase, at the B97-D3 level of theory[3] using the Becke-Johnson damping and the three-body term options (RI-B97-D3(BJ)-abc) with a def2-TZVPPD basis set,[4, 5] using the *jobex* module of the Turbomole package.[6] The resolution-of-identity (RI) approximation[7] and the associated auxiliary basis[8, 9] were also used. The normal modes, numerical harmonic frequencies and IR absorption strengths were calculated at the same level of theory using the *numforce* module. Gibbs energies were obtained at both 0 and 300 K, using the *freeh* module of Turbomole. A series of benchmark studies[10, 11] have shown that this level of theory provides a satisfactory agreement between gas phase populations and theoretical predictions.

Structures, energetics and frequencies in solution were carried out using the Conductor-like Screening Model approximation (COSMO),[12] available in the Turbomole package, where the solvent is modelled by a dielectric continuum of permittivity $\varepsilon = 4.81$ (for chloroform).

For comparison with experimental spectra in the NH stretch region, the harmonic frequencies were scaled by factors previously determined to provide a fair agreement with experiment (typical precision of *c.* 20 cm^{-1}) for compound **1**, namely 0.9780 and 0.9685, for gas phase and solution respectively.[1, 13] In the amide I and II regions, the harmonic frequencies were also scaled. In the absence of a previously determined scaling factor, the factor used for both regions was adjusted (1.014 for the chloroform solution) for the extended conformer (5^{O}-7^{δ} (*g*)) to best fit the frequencies of the most intense bands of the *trans* sulfoxide compound **2**.

2.2 Conformational landscapes

The conformational landscapes of compounds **2**-**4** (Figure S2.1) were similar to those found previously for compound **1**, consisting of three families: C5, C7, and δ conformations. Respectively, these structures are based on: i) an extended backbone (labelled 5 for short; yellow regions in Figure S2.1) stabilized by an intra-residue C5 and an inter-residue H-bond (C6γ or C7δ), implicating NH^1 and NH^2 respectively; ii) a folded backbone built around a C7 H-bond (7 for short; green regions) formed by NH^2; and iii) a partly folded structure (labelled δ for short; violet regions) wherein the amide NH^2 is oriented towards the preceding nitrogen in a π-amide (π_{am}) interaction.

In the gas phase (Figure S2.1), for compound **2** the 5 family was found to be the most stable, by typically 15 $kJ \cdot mol^{-1}$; this was also the case for compound **4**, although to a significantly lesser extent. For compound **3**, competition was clearly in evidence, with representatives of all three families appearing within 5 $kJ \cdot mol^{-1}$; the 5 family was slightly less stable.

In solution (Figure S2.1), a similar trend was observed, but with a significant stabilization of the δ families relative to the others, leading *in fine* to a predominant (but not exclusive) 5 form in **2**, to a prominent δ form for **3**, and to a rather open competition for **4** where δ, 7 and 5 forms are all expected.

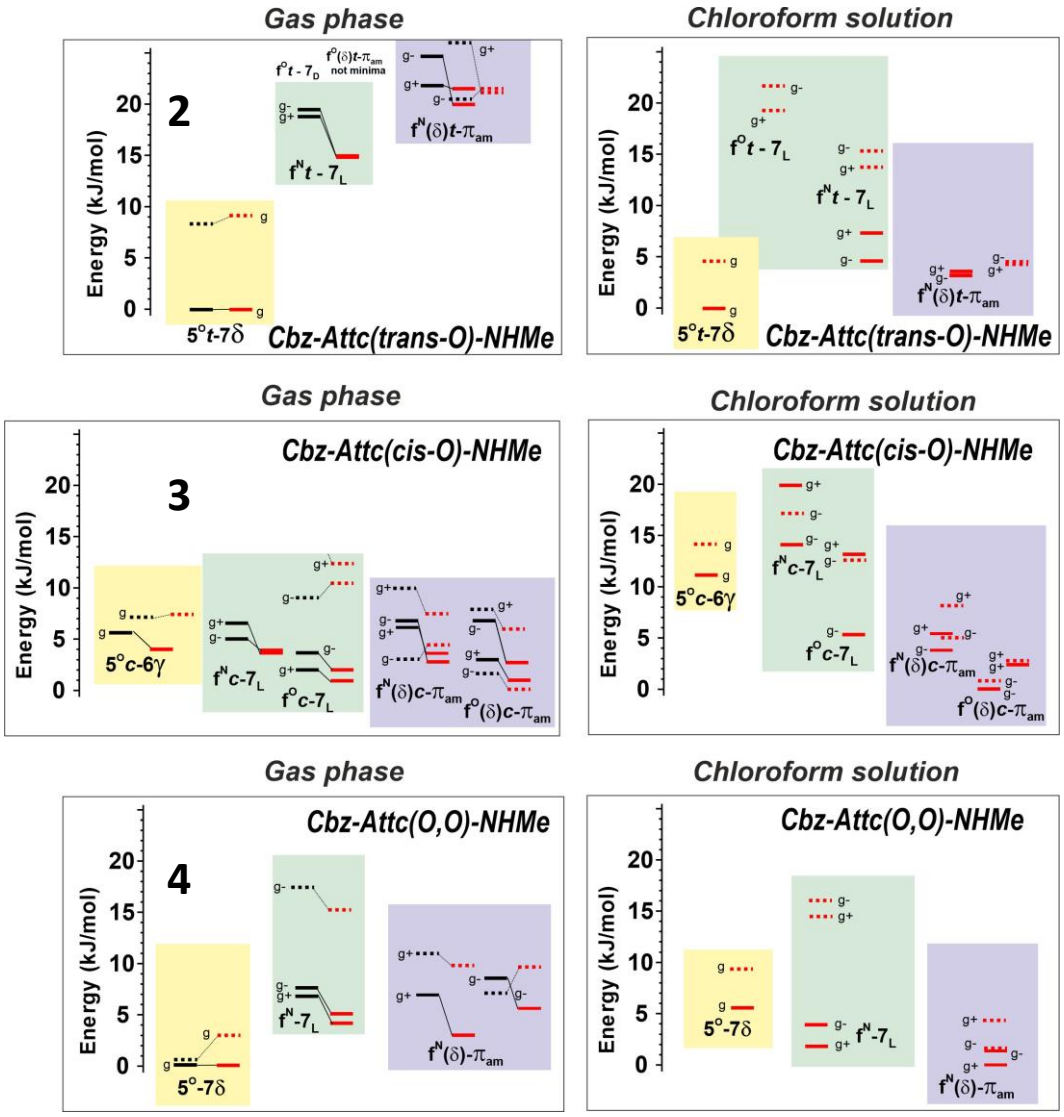


Figure S2.1. Energetic landscapes of the three backbone families of compounds **2**-**4** in the gas phase (left panel) and in chloroform solution (right panel). In the gas phase, two temperatures were considered: 300 K (red levels) in order to provide insight into the conformer populations in the supersonic expansions (usually well described by such high temperature) and 0 K (black levels) for which the energetics are close to the final conditions of the supersonic expansion and thus relevant to conformational relaxation phenomena within each family.[1, 14-16] For chloroform solution, a temperature of 300 K was assumed. Conformations are described according to the H-bonding status of their NH^1 and NH^2 moieties. The *t* and *c* labels in **2** and **3** refer to the *trans* and *cis* sulfoxide conformations, respectively. The label *g*+/– stands for the *gauche*+/– orientation of the Cbz moiety and the N/O superscripts indicate the puckering of the heterocyclic ring, towards the closest backbone atom, N or O. For the sake of simplification, the *trans* Cbz orientations were disregarded since they are usually slightly less stable than their *gauche* counterparts.[1, 14] In all cases, both *trans* and *cis* configurations of the carbamate were considered and the corresponding conformers are indicated by a full line and a dotted line, respectively.

2.3 Structures of the most stable forms of each family

The gas phase and solution structures of the most stable forms of each backbone family are depicted in Figures S2.2 and S2.3 respectively.

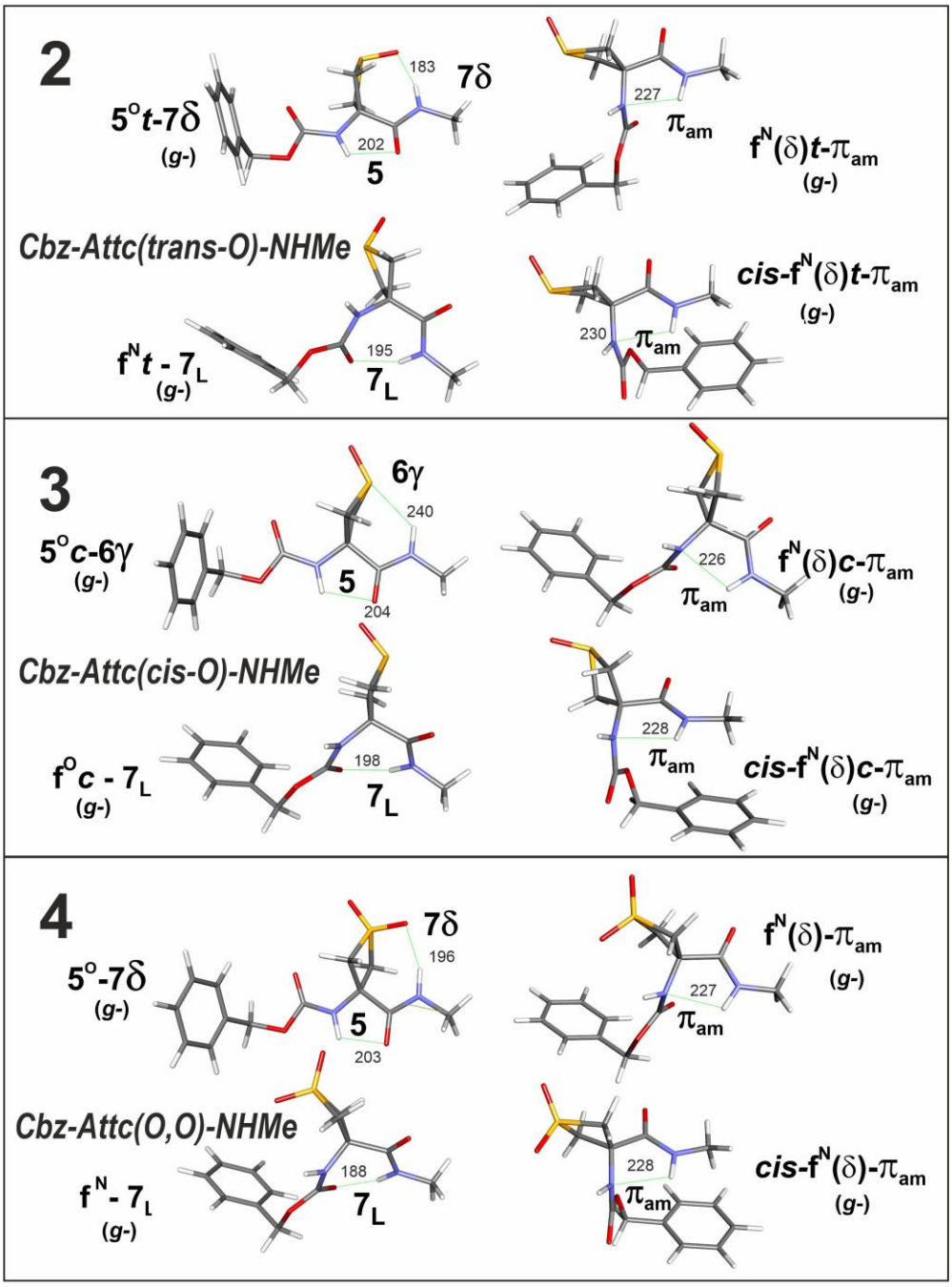


Figure S2.2. Calculated structures of the most stable forms of each backbone family of compounds **2-4** (from top to bottom) in the gas phase, illustrated with a *gauche*–orientation of the Cbz moiety. Conformations are labelled according to the H-bonding status of their NH^1 and NH^2 moieties. The *t* and *c* labels in **2** and **3** refer to the *trans* and *cis* sulfoxide conformations respectively. N/O superscripts indicate the puckering of the heterocyclic ring towards the closest backbone atom, N or O. Unless indicated by a *cis* prefix, all the configurations shown exhibit a *trans* carbamate. H-bonding distances are indicated in pm.

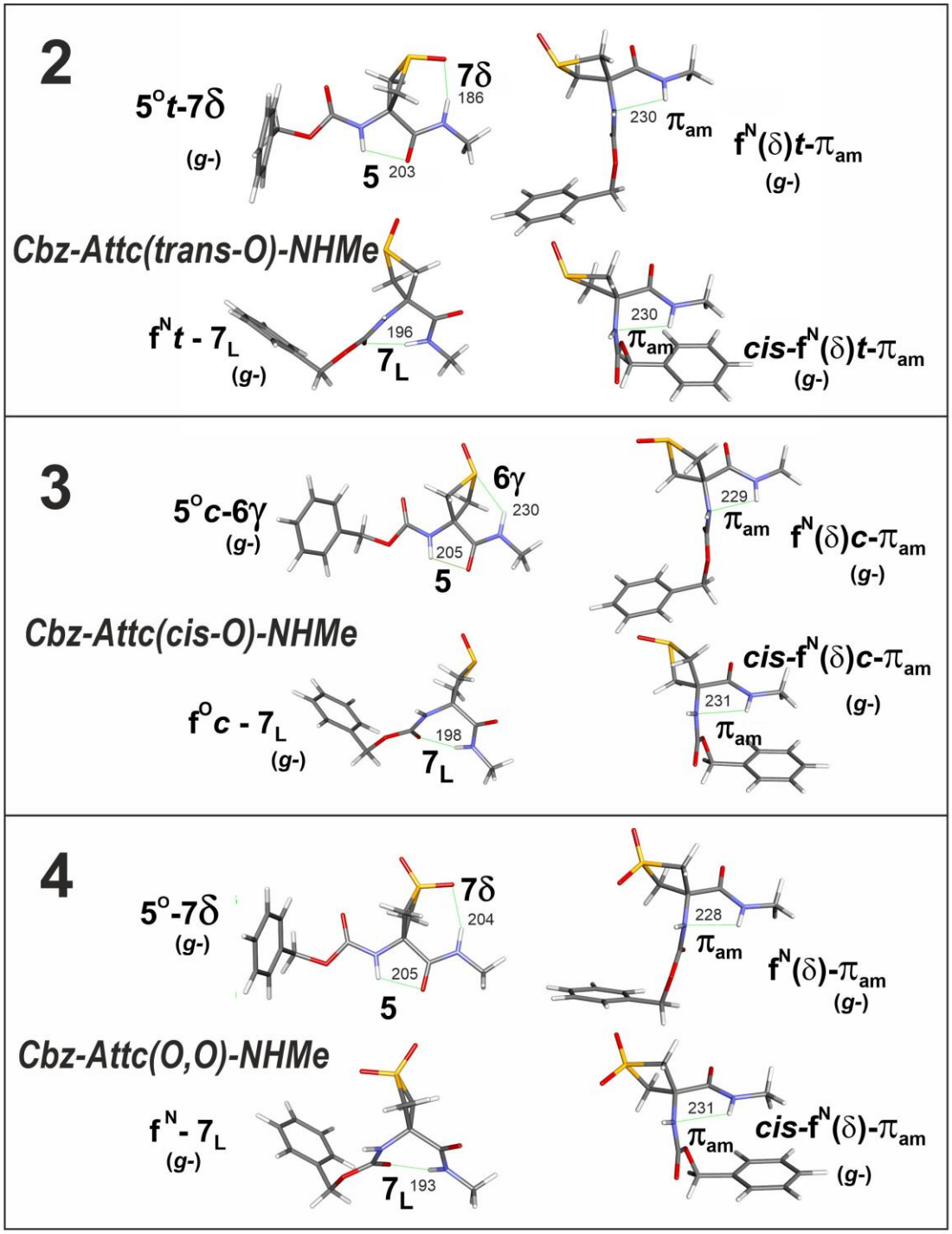


Figure S2.3. Calculated structures of the most stable forms of each backbone family of compounds **2**-**4** (from top to bottom) in a chloroform solution, illustrated with a *gauche*– orientation of the Cbz moiety. Same caption details as for Figure S2.2.

3. Solution state spectroscopic analysis

3.1. Infrared spectra

Solution state infrared spectra of compounds **1-4** were recorded at 295 K on a Fourier-transform Perkin Elmer Spectrum Two spectrometer, using 5 mM solutions in $CHCl_3$ held in Omni-cell Specac 1 mm path-length NaCl plates. The absorbance bands in the N-H stretch (amide A) region are shown in Figure S3.1.

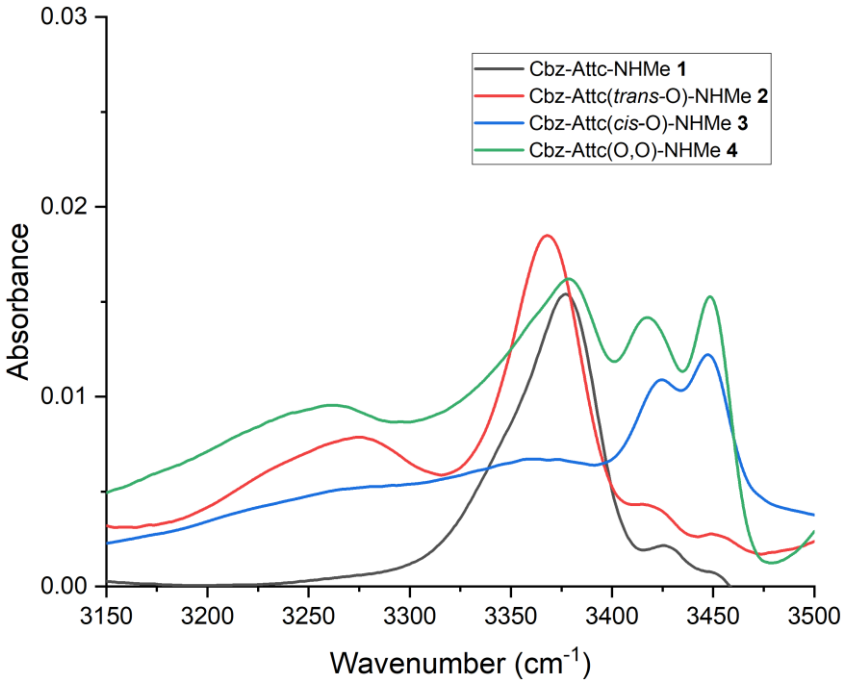
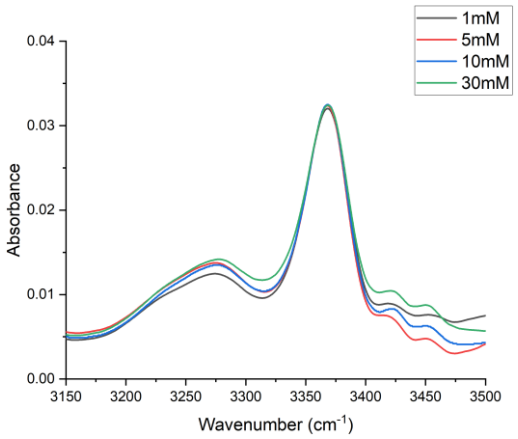


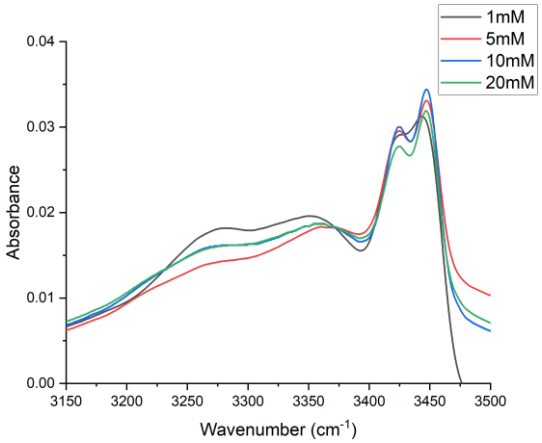
Figure S3.1. Solution state infrared spectra of compounds **1**-**4** (5 mM in chloroform).

It was previously shown that no concentration effects were in evidence in the IR spectrum of compound **1** in chloroform solution (range 1-30 mM).[14] Likewise, the IR spectra of **2** (range 1-30 mM) and of **3** and **4** (range 1-20 mM) showed no concentration effects (Figure S3.2). The upper limit of the concentration range for the latter two compounds represents their solubility limit in chloroform.

Cbz-Attc(*trans*-O)-NHMe **2**



Cbz-Attc(*cis*-O)-NHMe **3**



Cbz-Attc(O,O)-NHMe **4**

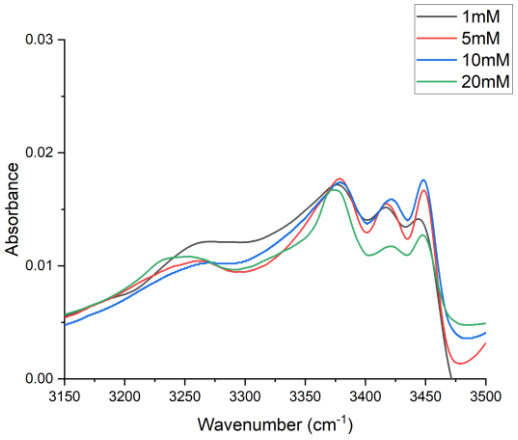


Figure S3.2. Solution state infrared spectra of compounds **2**-**4** at different concentrations.

3.2. ^{1}H NMR spectra

The solution-state behavior of compounds **2-4** in chloroform was also examined by ^{1}H NMR spectroscopy. In *trans*-sulfoxide **2** the carbamate NH^1 and amide NH^2 signals both appeared at low field, remarkably so for the latter (δ = 6.69 and 8.78 ppm, respectively) (Figure S3.3). Titration with DMSO-d_6 induced a moderate downfield shift for NH^1 and an upfield shift for NH^2 ($\Delta\delta$ = 0.76 and −1.62 ppm respectively) (Figure S3.4). In contrast, for both **3** and **4** the NH signals appeared at higher field (Figure S3.3) and titration with DMSO-d_6 induced a much larger downfield shift for NH^1 ($\Delta\delta$ = 2.14 and 2.22 ppm for **3** and **4** respectively) and a low downfield shift for NH^2 ($\Delta\delta$ = 0.37 and 0.12 ppm for **3** and **4** respectively) (Figure S3.4). These data are in good agreement with a predominant C5/C7δ conformer for compound **2** and a mixed conformational landscape for **3** and **4** in which folded C7 and π-amide interactions make significant contributions. The literature DMSO-d_6 titration data[1] for compound **1** (NH^1: $\Delta\delta$ = 0.60; NH^2: $\Delta\delta$ = –0.68) are comparable with those of compound **2** suggesting similar H-bonding status (*i.e.* both NH^1 and NH^2 are H-bonded).

Further evidence was obtained from qualitative NOESY experiments (Figure S3.5). In compound **2**, a strong correlation was observed between NH^2 and the *syn* protons of the 4-membered ring, as expected for a C5/C7δ conformer. In contrast, the corresponding correlation in compounds **3** and **4** was much weaker, whereas a strong correlation between NH^1 and the *syn* protons was in evidence, in agreement with folded conformations for these two compounds.

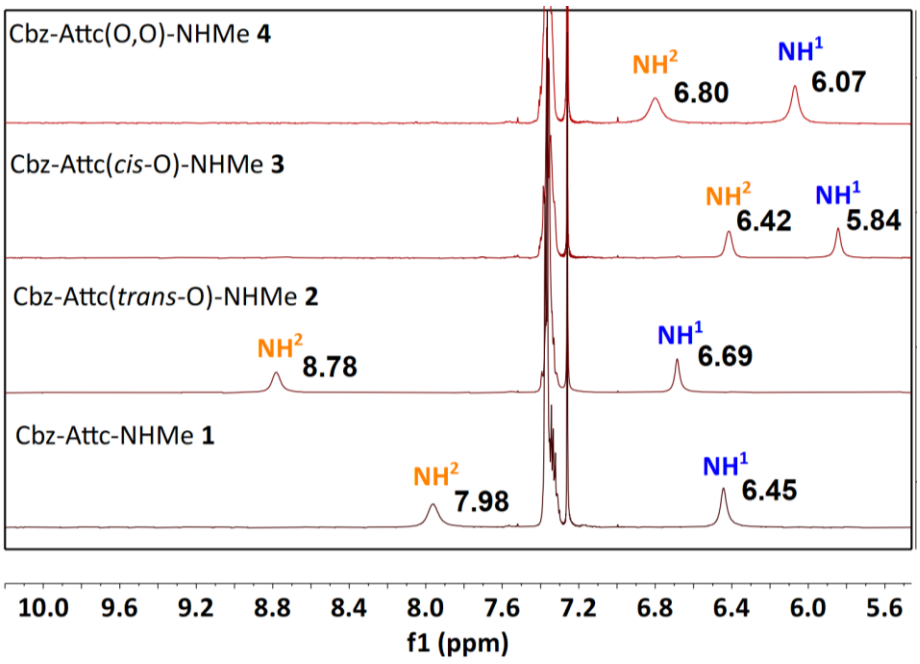
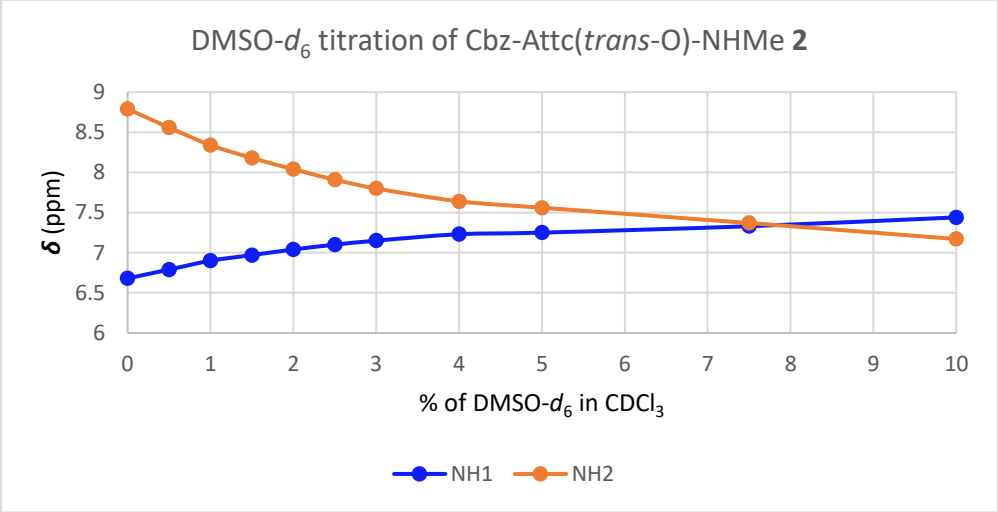


Figure S3.3. Expansion of the ^{1}H NMR spectral window where the NH proton signals of compounds **1-4** are located. Spectra were recorded at 300 K on a Bruker 400 MHz spectrometer. Solutions were 20 mM in $CDCl_3$.

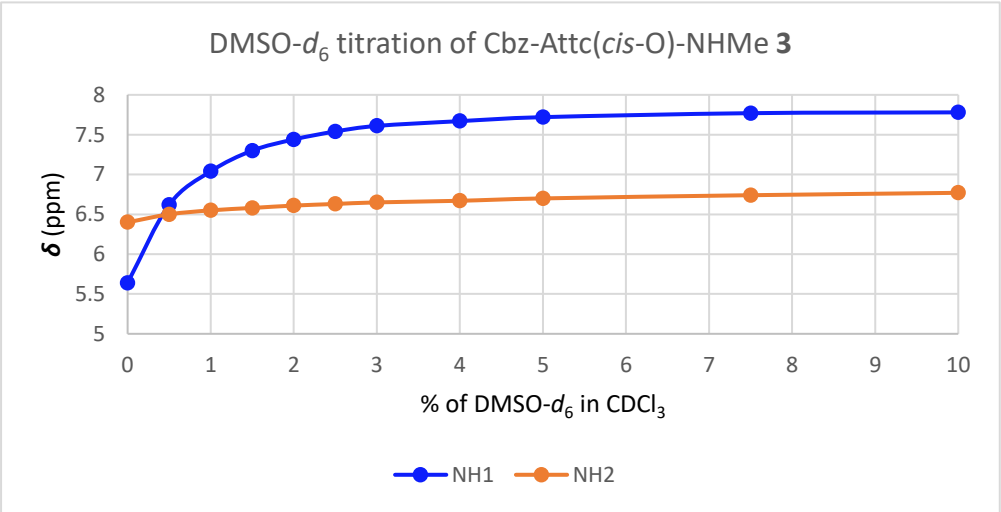
Cbz-Attc(*trans*-O)-NHMe **2**

DMSO-d_6 (%, v/v)	0	0.5	1	1.5	2	2.5	3	4	5	7.5	10	$\Delta\delta$
δNH^1	6.68	6.79	6.90	6.97	7.04	7.10	7.15	7.23	7.25	7.33	7.44	0.76
δNH^2	8.79	8.56	8.34	8.18	8.04	7.91	7.80	7.64	7.56	7.37	7.17	−1.62



Cbz-Attc(*cis*-O)-NHMe **3**

DMSO-d_6 (%, v/v)	0	0.5	1	1.5	2	2.5	3	4	5	7.5	10	$\Delta\delta$
δNH^1	5.64	6.62	7.04	7.30	7.44	7.54	7.61	7.67	7.72	7.77	7.78	2.14
δNH^2	6.40	6.50	6.55	6.58	6.61	6.63	6.65	6.67	6.70	6.74	6.77	0.37



Cbz-Attc(O,O)-NHMe **4**

DMSO-d_6 (%, v/v)	0	0.5	1	1.5	2	2.5	3	4	5	7.5	10	$\Delta\delta$
δNH^1	6.02	7.48	7.78	7.95	8.07	8.14	8.19	8.23	8.25	8.25	8.24	2.22
δNH^2	6.80	6.76	6.76	6.76	6.77	6.78	6.79	6.81	6.83	6.88	6.92	0.12

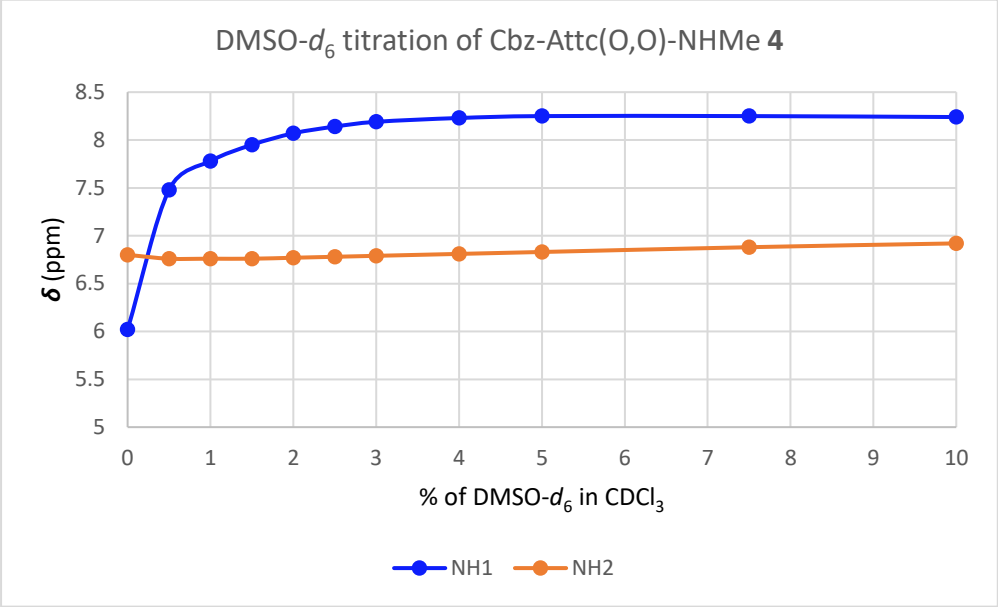
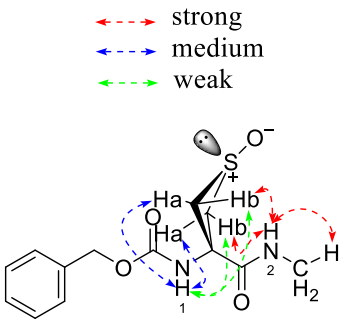
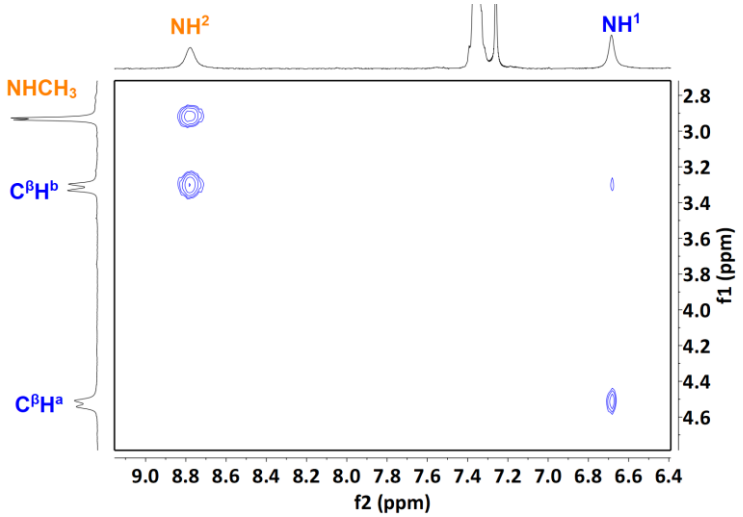


Figure S3.4. DMSO-d_6 titrations of compounds **2-4**. ^{1}H spectra were recorded at 300 K on a Bruker 400 MHz spectrometer. Samples were dissolved in $CDCl_3$ (400 µL) to give solutions of concentration 5 mM. Aliquots of DMSO-d_6 (6 × 2 µL, 2 × 4 µL, 2 × 10 µL) were added successively to the NMR tube, ^{1}H spectra were recorded after each addition. For each compound, the change in chemical shift ($\Delta\delta$) of each NH signal is given for 10% added DMSO-d_6.

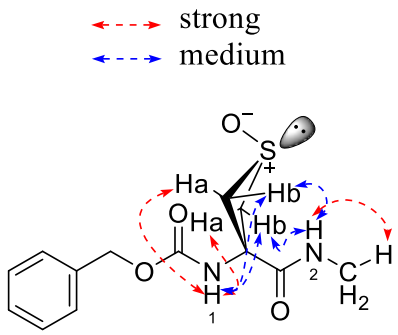
Cbz-Attc(*trans*-O)-NHMe **2**

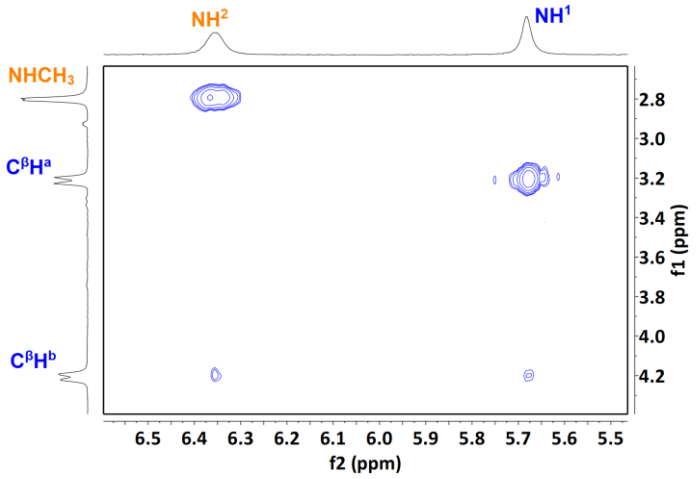




NH	NOE correlations	Intensity of NOE correlations
NH^1	$C^\beta H^a$	medium
	$C^\beta H^b$	weak
NH^2	$NHCH_3$	strong
	$C^\beta H^b$	strong

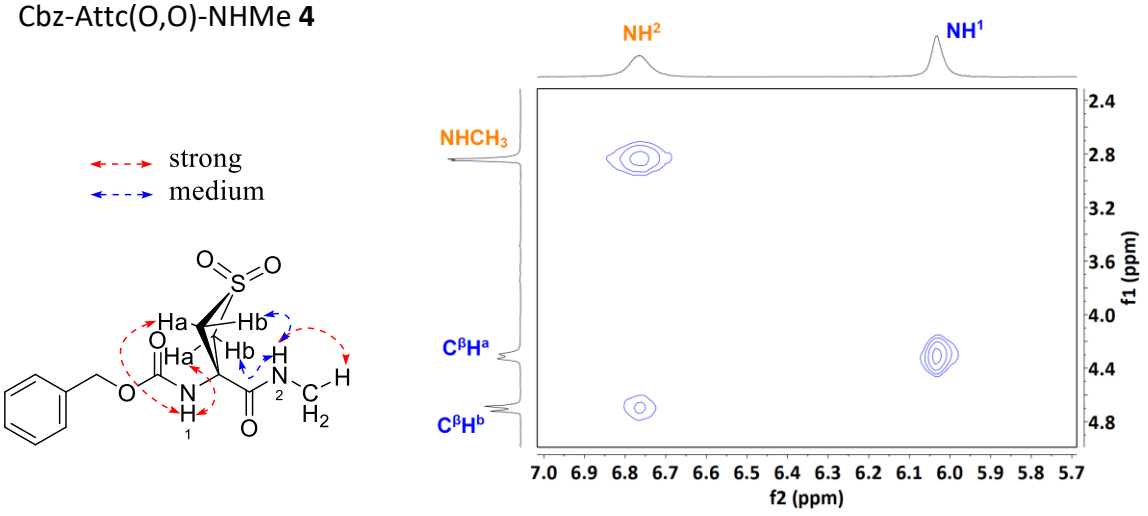
Cbz-Attc(*cis*-O)-NHMe **3**





NH	NOE correlations	Intensity of NOE correlations
NH^1	$C^\beta H^a$	strong
	$C^\beta H^b$	medium
NH^2	$NHCH_3$	strong
	$C^\beta H^b$	medium

Cbz-Attc(O,O)-NHMe **4**



NH	NOE correlations	Intensity of NOE correlations
NH^1	$C^{\beta}H^{a}$	strong
NH^2	$NHCH_3$	strong
	$C^{\beta}H^{b}$	medium

Figure S3.5. NOESY correlation plots of compounds **2**-**4**. Spectra were recorded at 300 K on a Bruker 400 MHz spectrometer. Samples were prepared in $CDCl_3$ at a concentration of 20 mM. The pulse sequence was *noesygpph* and the mixing time was 600 ms. The experiments were performed by collecting 2048 points in f2 and 256 points in f1.

4. Gas phase spectroscopic analysis

4.1. Principle of the experiment

The characterization of the most stable forms of the studied compounds in the gas phase was carried out using laser spectroscopies combined with quantum chemistry. Details of the experimental set-up and of the procedures used are published elsewhere.[15, 17] Vaporisation of the solid sample was carried out in vacuum using laser desorption. The resulting isolated molecules were subsequently cooled down in a pulsed supersonic expansion, operating with a mixture 70:30 of Ne:He gases, and then photo-excited and photo-ionized by interaction with a UV laser beam. The resulting ions were detected in a time-of-flight mass spectrometer, which enabled mass selection of the ion signals collected. Resonant two-photon ionization (R2PI) spectra in the region of the near UV absorption of the phenyl ring were recorded by collecting this signal as a function of the laser light wavenumber. The low temperature achieved (typically 10 K) for the translational and rotational degrees of freedom enabled conformer resolution due to the sensitivity of the phenyl ring absorption to the molecular conformation. Taking adavantage of this ability to detect selectively conformers one at a time, IR spectroscopy of each detected conformer was recorded in the NH stretch region, using the IR/UV technique. Briefly, an IR laser was shone on the molecular jet. When in resonance with a molecular transition, it depopulated the vibrational ground state of the molecules. This depopulation was subsequently detected by the UV laser, tuned on a transition of one conformer. Scanning the IR laser enabled recording conformer-selective IR spectra of each detected conformer. Comparison with relevant quantum chemistry simulations of the IR spectra enabled us to assign the detected conformers.

4.2. Results

The UV spectra of compounds **2**-**4** are shown on the left panel of Figure S4.1, together with that of compound **1** for reference. For **1** and **4**, pure samples were used. For **2** and **3**, investigated were conducted with two mixed samples, with *trans*:*cis* ratios of 3:2 and 1:20.

Compound **4** gave rise to a unique conformer, responsible for UV bands labelled A, A1 and A2, since these three bands gave rise to the same IR spectrum. The IR spectrum (Figure S4.1 right panel) was composed of two NH stretch bands, located at 3357 and 3393 cm^{-1}, and bore strong similarity with the spectrum of the main conformer of **1**. The fair agreement with the calculated spectrum of the most stable form found theoretically enabled us to assign this conformer to an extended backbone C5 conformation accompanied by a C7δ N–H···O=S inter-residue H-bond. This is supported by the previous assignment of the main form of compound **1** to an

extended backbone form stabilized by a C6γ H-bond (the so-called C5-C6γ motif).[1] Satisfactorily, the homogeneity of the UV transition frequencies is noticeable, together with IR transitions close to 3400 cm^{-1} (yellow region in Figure S4.1), diagnostic of a C5 H-bond in α,α-disubstituted α-aminoacid derivatives.[1, 14]

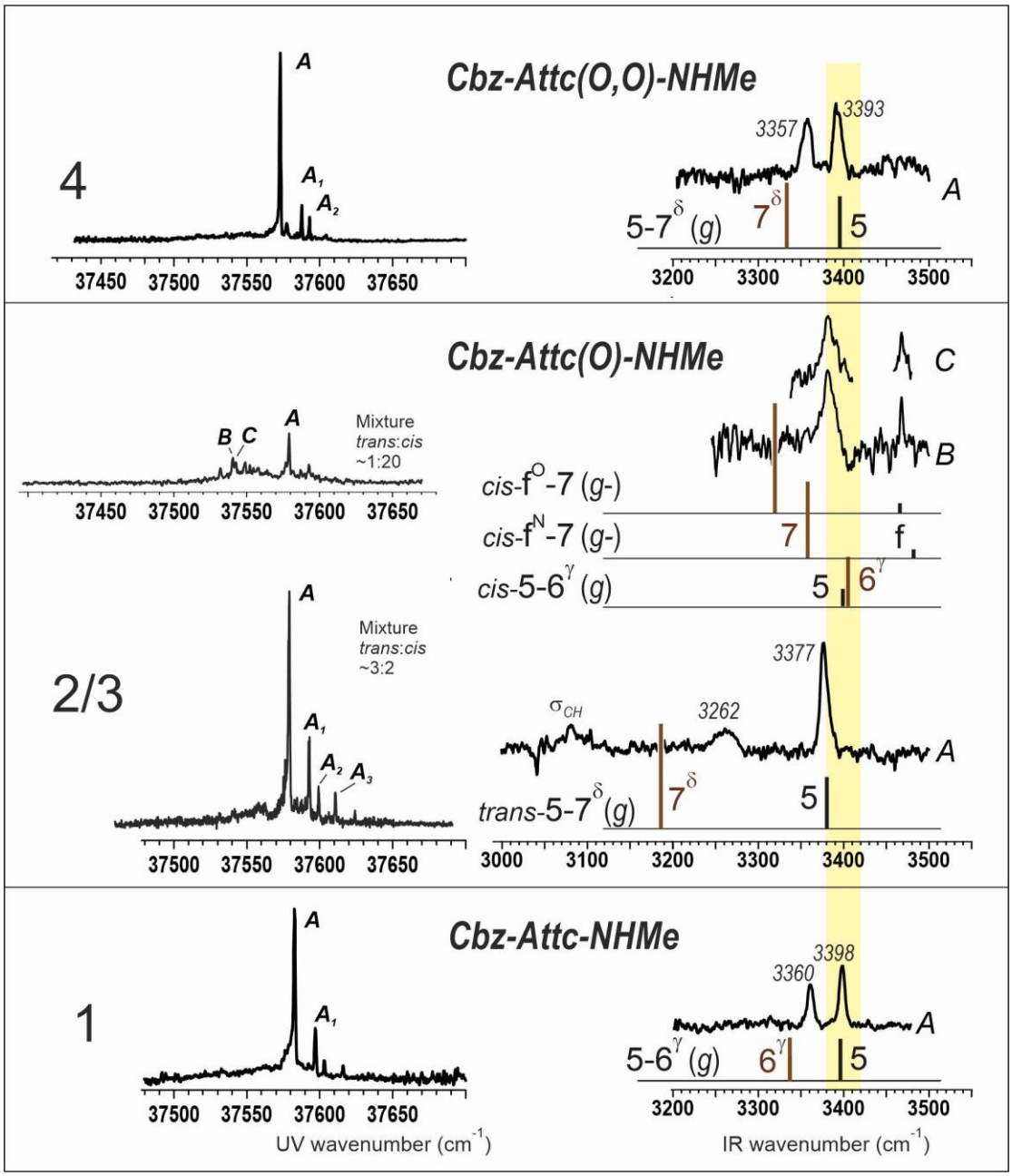


Figure S4.1. *left*: UV spectra for vaporised samples of mixtures of compounds **2** and **3** (*center panel*) or of compound **4** (*top panel*), obtained by mass-selected R2PI spectra at the origin of the first electronic ππ* transition of the phenyl ring in the near UV. *right*: Conformer-selective NH stretch IR spectra recorded by IR/UV spectroscopy using the main UV bands as conformer-selective probes, compared with relevant simulated spectra for the gas phase most stable forms of the compounds considered. For comparison, the corresponding spectra of compound **1**, adapted from Ref. 1, are also given (*bottom panel*).

Experiments on the **2**/**3** *trans*:*cis* mixtures required more attention. In the first experiment, conducted on a sample with a *trans*:*cis* ratio of 3:2, the IR spectra obtained from IR/UV spectroscopy using the four main UV bands observed (A, A1-3) were all identical (not shown), demonstrating the detection of a unique conformer. The IR spectrum was composed of a narrow band at 3377 cm^{-1} accompanied by a broader and much more red-shifted one at 3262 cm^{-1}, both being indicative of H-bonded NH groups. This observation excluded δ and 7 conformations, which both exhibit at least one free (or nearly free) NH moiety, and suggested extended backbones with 5 conformations. The 3262 cm^{-1} band exhibits the most red-shifted band of a H-bond in these systems,[1, 14] suggesting the involvement of a strong C7δ H-bond as expected for a *trans* sulfoxide C5-C7δ conformation. Comparison with simulated IR spectra of the most stable 5 conformer structures of both *trans* and *cis* sulfoxides (Figure S4.1) supports the assignment to the *trans* form: the C5 band is well reproduced, arising at slightly lower frequency than in other thiacyclic α-aminoacid derivatives, in agreement with the shorter C5 H-bond in this conformation. The red shift of the C7δ H-bond is found to be overestimated, which is a trend already encountered for the C6γ bonds in other thiacyclic α-aminoacid derivatives,[1, 14] suggesting a limitation in the general scaling factors presently used in the description of the NH···X H-bonds, *i.e.*, without distinguishing the nature of the acceptor.

It should be underlined that the alternative assignment, to the C5-C6γ conformation of a *cis* system, is disqualified by the too-weak C6γ H-bond of this form, as evidenced by the comparison with the theoretical IR spectrum (Figure S4.1). The unambiguous assignment of the unique observed conformation to the C5-C7δ backbone of the most stable form of the *trans* sulfoxide **2** in the gas phase (see Section S2) raises the question of detection of compound **3**, which contributed to 40 % of the population. For reasons that are unclear, no conformations of compound **3** were detected. This was confirmed by a second experiment conducted on a sample with a *trans*:*cis* ratio of 1:20 mixture, in which **3** represented 95 % of the population. The signal of the UV spectrum was weak; the most abundant signal (band A) was that of the main conformation of **2**. Only weak bands, labelled B and C, in the UV spectrum could be assigned to **3**, since these bands were not present in the UV spectrum of the first mixture (with ratio 3:2). The IR spectra obtained from the B and C bands were found to exhibit a single H-bonded absorption together with one nearly free NH band, in the 3470 cm^{-1} region, which reasonably matched C7-type conformations. Due to the expected stability of these structures, according to the gas phase landscapes calculated for **3** (see Section S2), the weakness of the signal confirms the poor detectivity of the resonant two-photon ionization scheme used in the case of **3**. This phenomenon is sometimes observed in cases marked by a short lifetime of the electronic state used as the intermediate level in the R2PI process,[18] but occurs very rarely for peptides. The only example we are aware of is the case of a β-turn conformation associated with a disulfide bridge in Ac-Cys-Phe-Cys-NH_2.[19]

5. References

1. Z. Imani, V. R. Mundlapati, G. Goldsztejn, V. Brenner, E. Gloaguen, R. Guillot, J. P. Baltaze, K. Le Barbu-Debus, S. Robin, A. Zehnacker, M. Mons and D. J. Aitken, *Chem. Sci.*, 2020, **11**, 9191.

2. D. Y. Liu, J. X. Bardaud, Z. Imani, S. Robin, E. Gloaguen, V. Brenner, D. J. Aitken and M. Mons, *Molecules*, 2023, **28**, 5046.

3. S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456.

4. D. Rappoport and F. Furche, *J. Chem. Phys.*, 2010, **133**, 134105.

5. A. Schafer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829.

6. Turbomole V7.2, **2017**, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, Turbomole GmbH, since 2007; available from http://www.turbomole.com.

7. M. Sierka, A. Hogekamp and R. Ahlrichs, *J. Chem. Phys.*, 2003, **118**, 9136.

8. K. Eichkorn, O. Treutler, H. Ohm, M. Haser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283.

9. K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theo. Chem. Acc.*, 1997, **97**, 119.

10. E. Gloaguen, B. de Courcy, J. P. Piquemal, J. Pilmé, O. Parisel, R. Pollet, H. S. Biswal, F. Piuzzi, B. Tardivel, M. Broquier and M. Mons, *J. Am. Chem. Soc.*, 2010, **132**, 11860.

11. Y. Loquais, E. Gloaguen, S. Habka, V. Vaquero-Vara, V. Brenner, B. Tardivel and M. Mons, *J. Phys. Chem. A*, 2015, **119**, 5932.

12. A. Klamt and G. Schuurmann, *J. Chem. Soc. Perkin Trans. 2*, 1993, 799.

13. Z. Imani, V. R. Mundlapati, V. Brenner, E. Gloaguen, K. Le Barbu-Debus, A. Zehnacker-Rentien, S. Robin, D. J. Aitken and M. Mons, *Chem. Commun.*, 2023, **59**, 1161.

14. V. R. Mundlapati, Z. Imani, V. C. D'Mello, V. Brenner, E. Gloaguen, J. P. Baltaze, S. Robin, M. Mons and D. J. Aitken, *Chem. Sci.*, 2021, **12**, 14826.

15. E. Gloaguen, M. Mons, K. Schwing and M. Gerhards, *Chem. Rev.*, 2020, **120**, 12490.

16. V. R. Mundlapati, Z. Imani, G. Goldsztejn, E. Gloaguen, V. Brenner, K. Le Barbu-Debus, A. Zehnacker-Rentien, J. P. Baltaze, S. Robin, M. Mons and D. J. Aitken, *Amino Acids*, 2021, **53**, 621.

17. E. Gloaguen, H. Valdes, F. Pagliarulo, R. Pollet, B. Tardivel, P. Hobza, F. Piuzzi and M. Mons, *J. Phys. Chem. A*, 2010, **114**, 2973.

18. M. Mons, F. Piuzzi, I. Dimicoli, L. Gorb and J. Leszczynski, *J. Phys. Chem. A*, 2006, **110**, 10921.

19. E. Gloaguen and M. Mons, *unpublished results*.