HAL
open science

A theoretical and experimental case study of the hydrogen bonding predilection of S-methylcysteine

Venkateswara Rao Mundlapati, Zeynab Imani, Gildas Goldsztejn, Eric Gloaguen, Valérie Brenner, Katia Le Barbu-Debus, Anne Zehnacker-Rentien, Jean-Pierre Baltaze, Sylvie Robin, Michel Mons, et al.

To cite this version:

Venkateswara Rao Mundlapati, Zeynab Imani, Gildas Goldsztejn, Eric Gloaguen, Valérie Brenner, et al.. A theoretical and experimental case study of the hydrogen bonding predilection of S-methylcysteine. Amino Acids, 2021, 53 (4), pp.621-633. 10.1007/s00726-021-02967-z . hal-03186714

HAL Id: hal-03186714

https://hal.science/hal-03186714

Submitted on 31 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A theoretical and experimental case study of the hydrogen bonding predilection of S-methylcysteine

Venkateswara Rao Mundlapati[1,a] · Zeynab Imani[2] · Gildas Goldsztejn[1] · Eric Gloaguen[1] · Valérie Brenner[1] · Katia Le Barbu-Debus[3] · Anne Zehnacker-Rentien[3] · Jean-Pierre Baltaze[2] · Sylvie Robin[2,4] · Michel Mons[1] · David J. Aitken[2]

Abstract

S-containing amino acids can lead to two types of local NH···S interactions which bridge backbone NH sites to the side chain to form either intra- or inter-residue H-bonds. The present work reports on the conformational preferences of *S*-methyl-L-cysteine, Cys(Me), using a variety of investigating tools, ranging from quantum chemistry simulations, gas phase UV and IR laser spectroscopy, and solution state IR and NMR spectroscopies, on model compounds comprising one or two Cys(Me) residues. We demonstrate that in gas phase and in low polarity solution, the C- and N-capped model compound for one Cys(Me) residue adopts a preferred C5-C6γ conformation which combines an intra-residue N−H···O=C backbone interaction (C5) and an inter-residue N−H···S interaction implicating the side-chain sulfur atom (C6γ). In contrast, the dominant conformation of the C- and N-capped model compound featuring two consecutive Cys(Me) residues is a regular type I β-turn. This structure is incompatible with concomitant C6γ interactions, which are no longer in evidence. Instead, C5γ interactions occur, that are fully consistent with the turn geometry and additionally stabilize the structure. Comparison with the thietane amino acid Attc, which exhibits a rigid cyclic side chain, pinpoints the significance of side chain flexibility for the specific conformational behavior of Cys(Me).

Keywords S-containing amino acids · NH···S H-bond · Conformation · Gas phase spectroscopy · Quantum Chemistry

Venkateswara Mundlapati and Zeynab Imani contributed equally to the work.

Electronic supplementary material The online version of this article (https ://doi.org/10.1007/….) contains supplementary material, which is available to authorized users.

✉ Michel Mons michel.mons@cea.fr
✉ David J. Aitken david.aitken@u-psud.fr

1 Laboratoire Interactions, Dynamiques et Lasers (LIDYL), CEA, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

2 Université Paris-Saclay, CNRS, Institut de Chimie Moléculaire et des Matériaux d'Orsay (ICMMO), 91405 Orsay, France

3 Université Paris-Saclay, CNRS, Institut des Sciences Moléculaires d'Orsay (ISMO), 91405 Orsay, France

4 Université de Paris, Faculté de Pharmacie, 75006 Paris, France

a Current Address: Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse (UPS), CNRS, CNES, 9 Avenue du Colonel Roche, 31028 Toulouse, France

Authors' ORCID

Venkateswara Rao Mundlapati	0000-0003-0559-9684
Zeynab Imani	0000-0003-0161-5106
Gildas Goldsztejn	0000-0002-9123-6419
Eric Gloaguen	0000-0002-1023-2791
Valérie Brenner	0000-0002-8004-1157
Katia Le Barbu-Debus	0000-0003-4019-4973
Anne Zehnacker	0000-0001-5540-0667
Jean-Pierre Baltaze	0000-0001-7271-8538
Sylvie Robin	0000-0003-3450-1412
Michel Mons	0000-0001-9930-831X
David J. Aitken	0000-0002-5164-6042

Introduction

Nature makes use of a number of low-molecular weight sulfur-containing compounds, including such diverse molecular structures as lenthionine, biotin, lipoic acid, thiamine, and of course the proteinogenic amino acids methionine and cysteine.(1977; 1980) A closely related sulfur-containing amino acid is *S*-methyl-L-cysteine, Cys(Me), which is not genetically coded but is formed by post-translational methylation of cysteine.(Maw 1982; Thompson 1967)

Cys(Me) is found in beans and legumes, both in its free amino acid form and as the dipeptide γ-glutamyl-S-methylcysteine. It is also one of the sulfur-containing amino acid constituents found in garlic, and it has been suggested that such organosulfur compounds may contribute to the antioxidant properties and health benefits of garlic as a foodstuff.(Liu et al. 2020; Zawistowski et al. 2018) It appears in some proteins, such as methyl-coenzyme M reductase (MCR) which catalyzes the final step in the biological production methane by methanogenic archaea.(Grabarse et al. 2000) A dipeptidyl boronic acid containing Cys(Me) was recently reported to be a proteasome inhibitor with promising anti-tumor activity,(Lei et al. 2019) while in industry, copolymers incorporating Cys(Me) are known to be efficient inhibitors of metal corrosion.(Mazumder 2019; Amin et al. 2010)

During the pioneering studies on the role of hydrogen bonding in determining the preferred conformations of peptides,(Toniolo 1980) Toniolo's group examined the secondary structures of the N- and C-capped oligomers Boc-[Cys(Me)]$_n$-OMe (n = 2-7).(Bonora et al. 1975; Palumbo et al. 1976; Toniolo et al. 1975) From a detailed analysis of solid state IR spectroscopic data, it was deduced that while the dimer was essentially disordered, a pleated β-sheet conformation stabilized by intermolecular H-bonds was present along with disordered forms for the trimer and the tetramer. The β-conformation was the preponderant structure for the pentamer to the heptamer (Palumbo et al. 1976). Solution state CD measurements in polar media suggested likewise that β-sheet conformers prevailed for n = 5-7 while ordered structures were less in evidence for the shorter oligomers (Toniolo et al. 1975). These data were in agreement with previous observations that polymers of Cys(Me), in contrast to polymers of methionine, are non-helical and adopt a β-structure. (Bloom et al. 1962; Fasman and Potter 1967; Komoto et al. 1974; Birdi and Fasman 1972) It was suggested that this may be the consequence of the presence of the sulfur atom in the γ-position of a Cys(Me) residue which might form a hydrogen bond with a backbone amide NH. Support for this hypothesis emerged from IR studies of dilute $CDCl_3$ solutions of Boc-[Cys(Me)]$_3$-OMe and Boc-[Cys(Me)]$_4$-OMe. While it was recognized that several more-or-less ordered conformers might co-exist for these small peptides, the frequencies of the amide A vibrations were apparently red-shifted compared to those of the corresponding derivatives of aliphatic amino acids, which was commensurate with the presence of N–H···S interactions (Palumbo et al. 1976).

Information on NH···S H-bonding interactions of small peptides including a residue bearing a sulfur atom in the side-chain γ-position has been obtained from gas phase studies.(Biswal et al. 2012; Yan et al. 2014; Alauddin et al. 2015; Goldsztejn et al. 2020a; Goldsztejn et al. 2020b) For example, the role of short-range C5γ intra-residue N–H···S bonds in stabilizing sequential γ-turns in Cys-containing peptides has been demonstrated (Yan et al. 2014; Alauddin et al. 2015; Goldsztejn et al. 2020a; Goldsztejn et al. 2020b) (Fig. 1). In a recent discovery from our consortium combining gas phase and solution phase studies, a thietane amino acid with a sulfur in the γ-position, Attc, was found to stabilize extended backbones at the expense of folded ones in short oligomers through the formation of short-range C6γ inter-residue N–H···S interactions (Fig. 1) (Imani et al 2020). This significant observation raises the question of whether such interactions might also prevail in short Cys(Me) oligomers. The aims of the present study were therefore to examine in detail the conformational behavior of the N- and C-capped derivatives Cbz-Cys(Me)-NHMe and Cbz-[Cys(Me)]$_2$-NHMe in the gas phase (where Cbz stands for the benzyloxycarbonyl group), aided by theoretical calculations, to correlate these findings with spectroscopic data obtained for the same derivatives in solution state, and to assess the extent to which Attc and Cys(Me) share their intrinsic folding proclivities (Fig. 2).

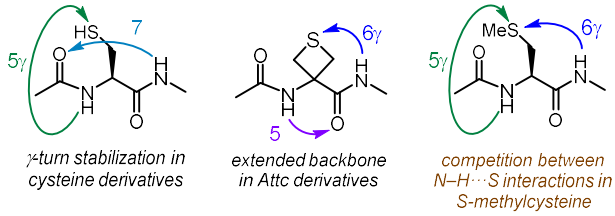

Fig. 1. (*left and centre*) Short range N–H···S interactions (in addition to other interactions) previously identified through gas phase studies of amino acids with a sulfur in the γ-position; (*right*) the corresponding interactions suggested for Cys(Me) which are probed in this work.

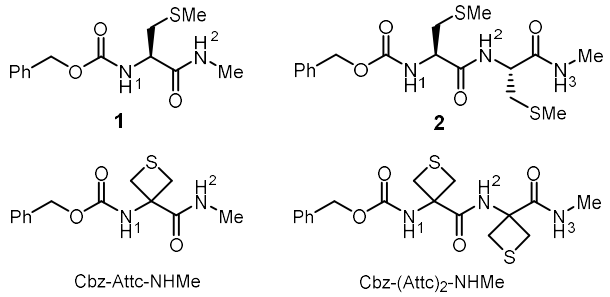

Fig. 2. Compounds examined and compared in this work.

In a context where competition between several H-bonding patterns can be anticipated for small peptides leading to conformational mixtures that are difficult to analyze in solution (such as the present study), gas phase conformational analysis appears as a valuable technique (Gloaguen and Mons 2015; Gloaguen et al. 2020). Molecules are brought into the gas phase from a solid sample using laser desorption and are entrained into a vacuum via supersonic expansion. After an early expansion stage, in which the stable backbone families at "high" temperature (typically room temperature, see (Gloaguen et al. 2010a) for example) are populated (Fig. 3), the vaporized molecules are rapidly cooled in the supersonic expansion through collisions with the carrier gas, with two outcomes: first, conformational relaxation occurs in each backbone family basin, eventually populating low energy conformers; second, an efficient rovibrational cooling, at a typical temperature of 10 K, allows vibrationally-resolved spectra to be recorded. Using laser techniques to combine IR and UV spectroscopies, conformer-selective vibrationally-resolved IR spectra can be obtained. The spectra can be assigned with the help of theoretical chemistry, including potential energy surface exploration at high temperatures and search for local minima, followed by quantum chemistry simulations of the IR spectra. This experimental technique requires an appropriate chromophore for UV detection, which is provided conveniently by the Cbz group of the derivatives studied in this work.

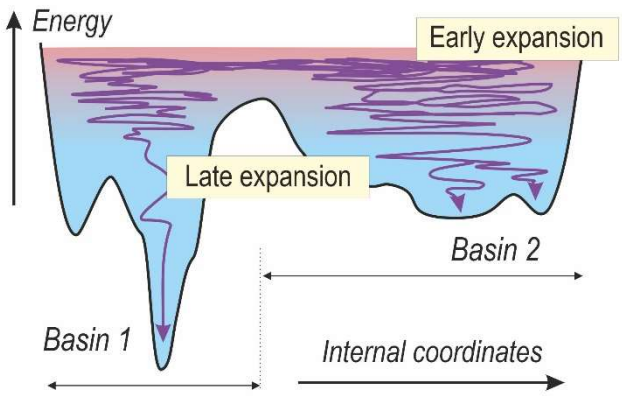

Fig. 3. Schematic representation of the conformational dynamics along the expansion. In the early expansion (typically room temperature) basins are populated according to the free energy of their components in this temperature range (red energetic region). During the expansion, the temperature goes down rapidly, freezing the conformational populations within the basins separated by high barriers, typically corresponding to different backbone and/or H-bonding arrangements (e.g. basins 1 and 2, assigned to different families). Within each basin, however, conformations are usually separated by lower barriers and interconversions are efficient enough for conformational relaxations to occur. The process takes place until the late expansion where the low temperature (blue region) eventually forbids interconversions, resulting in a final distribution, where only the lowest conformations are populated (e.g. basin 2, typically corresponding to Cbz-rotamers; see text). The low temperature eventually achieved in the expansion enables detection of resolved UV or IR spectroscopic features.

Materials & Methods

Synthesis

Compounds **1** and **2** were prepared from commercial *S*-methyl-L-cysteine using standard solution phase procedures. Full details of the syntheses and routine compound characterization are given in the Supplementary Material document (Section S1). Cbz-Attc-NHMe and Cbz-$(Attc)_2$-NHMe were prepared as described previously (Imani et al. 2020).

Theoretical Chemistry

The strategy of conformational assignment from gas phase spectroscopy stems from a detailed comparison of conformation- and vibration-resolved experimental data obtained in a supersonic expansion, i.e. at a low temperature (T~ 10 K) with the theoretical spectral signatures of the individual conformations populated in the expansion. As described above, these populations widely result from the tempering of the initial "hot" conformational distribution of the gas phase experiment due to the laser desorption process (Fig. 3). Our analysis is thus based on an assessment of such a high temperature conformational distribution, from conformational landscapes at both 300 K and 0 K, to determine among the populated structures at ambient temperature, those that can be populated in the expansion. This is followed with a "spectroscopic grade" simulation of vibrational spectra at low temperature (0 K), for a spectral identification of the structures that may be populated at the end of the supersonic expansion.

An extensive and detailed picture of the potential energy landscape of each compound in the gas phase was obtained for compounds **1** and **2**, through a dual step procedure combining a force field exploration and quantum chemistry structure optimizations. The exploration was done with the OPLS-2005 force field using the Monte-Carlo Multiple Minima procedure in the MacroModel suite (Macromodel Schrödinger Release 2019-3). Subsequent geometry optimizations were done using the Density Functional Theory (DFT) at the RI-B97-D3(BJ)-abc/def2-TZVPPD level of theory (Eichkorn et al. 1997; Grimme et al. 2010; Rappoport and Furche 2010) with the Turbomole 7.2 package (2017) using parameters (gridsize m3; SCF convergence threshold 10^{-8} a.u.; gradient norm convergence threshold 10^{-5} a.u.), which were successfully used previously.

The nomenclature used to describe the conformations is based on a sequential analysis of the H-bonding status of the NH groups along the peptide chain. When a hydrogen bond occurs, a number indicates the number of atoms involved in the ring formed by the H-bond between the NH being considered and its H-bond acceptor (carbonyl O or S atom). A subscript to the number of an N–H···O=C H-bond (*e.g.* 7_D) stands for the configurational stereochemistry type for a chiral L or D conformation. An NH group free from any hydrogen bonding is designated by the label "f".

For each backbone conformation, the Cbz group gives rise to three rotamers, a phenomenon that we refer to as Cbz-rotamerism. They are designated according to their C_{sp2}-C_{sp3}-O-C(=O) dihedral φ, namely *gauche*+ (g^+; φ = +60°), *gauche*– (g^-; φ = –60°) and *trans* (*t*; φ = ± 180°).

Subsequently, for each structure resulting from an optimized geometry at the quantum chemistry level, a frequency calculation was performed at the same level of theory within the harmonic approximation. These calculations enabled us to check that the conformations obtained corresponded to "true" minima and the harmonic frequencies derived were then scaled in order to account for anharmonicity and potential deficiencies of the level of theory used, leading to theoretical data directly reconciled with experimental data. Previous work on peptides in which computationally calculated conformational landscapes were found to give a good match with the structures observed experimentally following a supersonic expansion (Alauddin et al. 2015; Gloaguen et al. 2010a; Gloaguen et al. 2020) suggests that the accuracy of the Gibbs energies is of the order of a few (2-4) kJ/mol.

In the NH stretching (amide A) region, a fixed scaling factor of 0.978 was applied, based on a fit of frequency data on a set of capped peptides studies in the gas phase (Gloaguen and Mons 2015). This strategy generally provides theoretical frequencies in good agreement with experimental frequencies (within 20 cm^{-1} for NH stretches in NH···O bonds (Gloaguen and Mons 2015) and typically within ~30 cm^{-1} for NH···S bonds (Biswal et al. 2012; Imani et al. 2020)). In the CO stretch (amide I) and NH bend (amide II) regions, the scaling factors used (1.004 and 1.008, respectively) were chosen to fit the corresponding spectra of Attc (Imani et al. 2020) or Cys derivatives (Goldsztejn et al. 2020a).

In addition, a Natural Bond Orbital (NBO) detailed analysis (Reed et al. 1988; Alabugin et al. 2011; Weinhold 2012; Alabugin et al. 2019) was performed for each conformation, in order to quantify the stabilizing role of electron delocalization through hyperconjugation and H-bonding effects. The analyses were carried out on the RI-B97-D3(BJ)-abc/def2-TZVPPD structures using the NBO module (Weinhold et al.) of the Gaussian 09 software (Frisch et al. 2016), first at the HF/TZVPP level followed by the MP2/TZVPP level, the latter having been shown to provide a good compromise between accuracy and computation times (Gloaguen et al. 2014). The donor and acceptor NBO occupancies considered in the following are those obtained at the MP2/TZVPP level whereas the HF level was used to calculate, for each donor NBO(i) and acceptor NBO(j), the E(2) stabilization energy through the second order perturbation theory:

$$E(2) = \Delta E_{ij} = q_i \frac{F(i,j)^2}{|\varepsilon_i - \varepsilon_j|}$$

where q_i is the donor orbital occupancy, F(i, j) is the off-diagonal NBO Fock matrix element and ε_i (resp. ε_j) is the diagonal element, *i.e.* the donor (acceptor) orbital energy; the threshold for considering the interactions as significant was 0.2 kJ mol^{-1}.

Gas phase techniques

Compounds **1** and **2** were introduced into the gas phase using laser-desorption coupled to a supersonic expansion, as described in detail previously (Gloaguen et al. 2010b). Laser-desorption was carried out downstream from a pulsed nozzle by sending the frequency-doubled output of a pulsed Nd:YAG laser (Minilite Continuum, 10 Hz, 0.1-3 mJ/pulse) on the surface of a mixed graphite:sample (4:1 molar ratio) 6 mm-diameter pellet through an optical fiber. The vaporized molecules were cooled by interaction with the supersonic molecular beam generated from the expansion of a carrier gas (30:70 Ne:He mixture) at 18 bar pressure through a 1 mm diameter nozzle, before entering the source of a time-of-flight mass spectrometer (TOF-MS), by passing through a skimmer located 18 mm downstream the nozzle. The output of a frequency-doubled dye laser (NarrowScan, Radiant Dyes) (~0.4-0.7 mJ in the UV) pumped by a frequency tripled Nd:YAG (Powerlite, Continuum) was then used to excite and ionize the molecules during the same laser pulse (since the ionization energy of the chosen UV chromophore, the Cbz group, is smaller than twice the UV photon energy). Ions thus generated were accelerated and sent to the field-free region of the TOF-MS and eventually collected on a microchannel plate-based detector, after reflection on an electrostatic mirror, designed to lengthen the TOF and then enhance the mass resolution. The mass-selected ion signal was collected as a function of the UV frequency, yielding a resonant-two-photon ionization (R2PI) spectrum of each compound in the origin region of the $\pi \rightarrow \pi^*$ electronic transition of the aromatic ring of the Cbz group. The low temperature of the final expansion enables recording of the spectra of cold species, characterized by narrow lines due to the origin band of each conformer present in the expansion, possibly accompanied by (usually weak) vibronic bands, associated to the excitation of low frequency modes (usually in the 10-100 cm^{-1} range from the origin), if the molecular structure changes between ground and excited state (Franck-Condon activity).

The exquisite sensitivity of the chromophore excitation frequency to its environment, and hence to the backbone structure, causes the spectral UV signatures of the several conformer to appear at distinct frequencies and give rise to well-identified UV bands. This feature enabled us to probe selectively each cold conformer, and thus to provide conformer-selective IR spectra, using the so-called IR/UV double resonance spectroscopy. This technique relies on the fact that an IR laser tuned on an IR transition of a conformer will depopulate its

vibrational ground state. Then, the ion signal generated by a UV laser tuned on a transition of this particular conformer will experience a depletion, which can be easily detected and recorded as a function of the IR laser frequency, yielding to a conformer-selective IR absorption spectrum. Gas phase conformer-specific IR spectra of compounds **1** and **2** were recorded using this technique (Gloaguen and Mons 2015; Gloaguen et al. 2010b). The IR beam was generated either directly from the idler output of a Nd:YAG-pumped (Continuum Surelite, 740 mJ/pulse at 1064 nm) OPO/OPA (Laservision), yielding typically 10 mJ in the 3200-3600 cm^{-1} amide A range, or from the output of a Difference Frequency Generation (DFG) stage, downstream the OPO, where OPA idler and signal were converted to 1300-1800 cm^{-1} radiation (1 mJ output) in order to probe the amide I and II regions.

The conformers observed in the expansion are identified from the comparison between the IR experimental spectra thus obtained and the theoretical spectra obtained for the most stable structures at high temperature (300 K). It should be noted that the three rotamers of the Cbz group, which provides the UV chromophore required for the IR/UV double resonance spectroscopy, share essentially the same amide backbone. Depending on the final temperature achieved and the barrier between them, several rotamers can be observed: they give rise to close lying UV bands and to very similar IR spectra since these latter do not significantly depend on Cbz-rotamerism. The whole population should then be considered as representative of the population of their backbone at high temperature.

Solution phase techniques

Fourier transform IR spectra were recorded at ambient temperature (*c* 295 K) for solutions in $CHCl_3$ (5 mM), held in a 1.0 mm path-length NaCl solution cell, with background solvent subtraction. No significant changes were observed upon sample dilution, indicating the intramolecular nature of any prevalent non-covalent interactions.

^{1}H NMR spectra were recorded at 273 K using standard pulse sequences on a 600 MHz spectrometer for solutions in $CDCl_3$ (20 mM). Titration experiments were carried out at 300 K on a 400 MHz spectrometer in $CDCl_3$ (5 mM) with solution volumes of 400 μL. Aliquots of DMSO-d_6 (6 × 2 μL, 2 × 4 μL, 2 × 10 μL) were added successively, each addition was followed by rapid agitation then re-recording of the spectra. The NOESY experiment was carried out at 273 K on a 600 MHz spectrometer using a solution of compound **2** in $CDCl_3$ (20 mM). The pulse sequence was noesygpph, and the mixing time was 600 ms. The experiment was performed by collecting 1024 points in f2 and 512 points in f1.

Results & Discussion

Cbz-Cys(Me)-NHMe 1

Quantum chemistry investigation

The conformational landscape of **1** was obtained from a systematic exploration of the potential energy surface first at the force field level, further refined at the quantum chemistry level using a DFT-D method. At 300 K (Fig. 4), the most stable backbone family, composed of 4 sub-families (1-a to 1-d, in Fig. 4), had an extended backbone, featuring a C5 intra-residue interaction implicating NH(1), accompanied by a weaker interaction NH⋯S, bridging the NH(2) group of the C-terminal protection to the sulfur in the side-chain γ-position, labelled C6γ. The several sub-families, each composed of the three Cbz rotamers, differed by the side chain and carbamate conformations (details given in Table S1); the 1-a sub-family being the most stable by more than 5 kJ/mol, was characterized by a specific [t, g+] arrangement of the side chain, which enabled the formation of a short H-bond, with a NH⋯S distance of 245 pm.

The next conformational families (2-4) in energetic ordering at 300 K were all higher in energy by at least 4 kJ/mol; most of them benefit from the presence of an intra-residue C5γ NH⋯S bond. Some of them exhibited a folded backbone, either featuring a γ-turn (sub-families 3-a, 3-b and 4-a, 4-b) or an α-type Ramachandran conformation, with nearly perpendicular peptide planes and stabilized by weak π_{am1} H-bonds involving NH(2) and the π cloud of the first amide (sub-families 2-a and 2-b).

Interestingly, the conformational landscape of **1** is comparable with Cbz-Attc-NHMe (Supplementary information of (Imani et al. 2020)), with the same families present but with a much narrower spread in energy between the several families. Such a landscape suggests that family 1-a should represent the main conformation populated, at least in the gas phase. Regarding the rotamer population, from the small energetic difference observed at 0 K (Fig. 4), the observation of several rotamers of family 1-a can be anticipated in the late expansion.

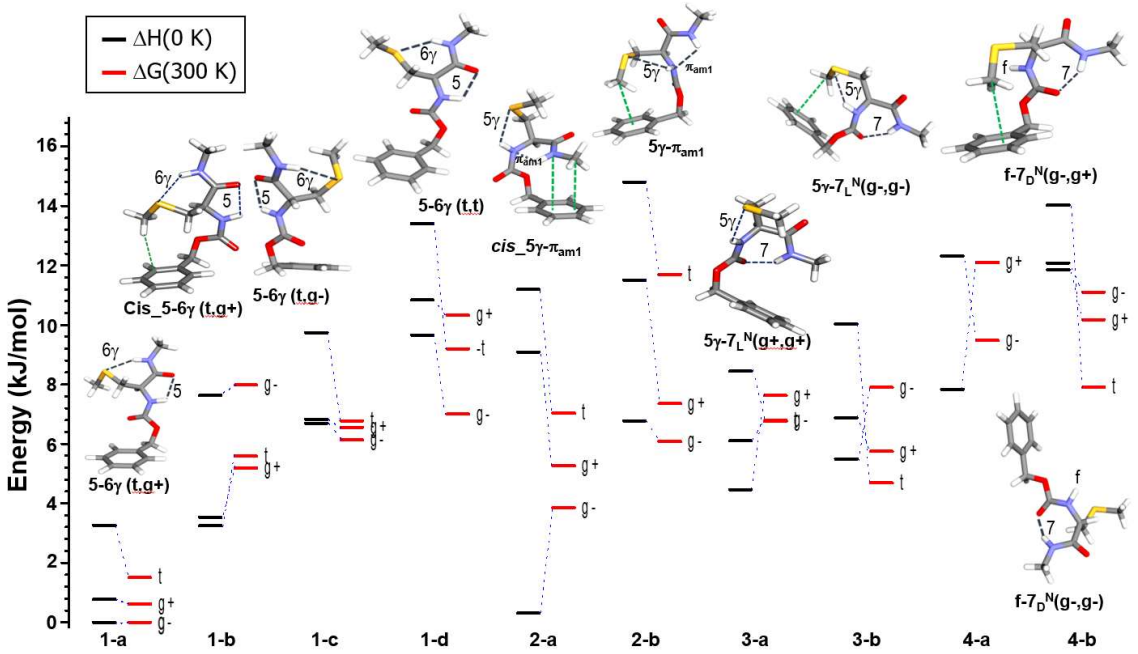

Fig. 4. Conformational landscape of **1** at RI-B97-D3(abc)/def2-TZVPPD level of theory, for two temperatures: 300 K (red lines) and 0 K (black). Intramolecular H-bonds and interactions stabilizing the structures are highlighted and labeled using short notations: 5, 7_L or 7_D, or π_{am1} for intra-backbone interactions, and 5γ and 6γ, for side chain – backbone H-bonds. The side chain orientation is indicated by the successive dihedrals (N-Cα-Cβ-S and Cα-Cβ-S-C_{Me}) given between brackets. Carbamate groups are in a *trans* conformation, unless indicated by a *cis* label. Only one Cbz rotamer is displayed for each sub-family.

Table 1. Theoretical structural backbones and H-bonding features (type, distances, NH stretch frequencies, E(2) NBO strength indicator) of the conformations observed for compounds **1** and **2** and related Attc compounds obtained at the RI-B97-D3(BJ)-abc/def2-TZVPPD level of theory. Additional structural and energetic data and comparison with experiment are provided in the Supplementary Material document (Sections S2.1 & S2.4).

Compound Conformer exp. obs'd	H-bonding type **Conformation sub-Family (Cbz rotamer)**	NH(i)	H-bond type(s)	NH(i)···X distance(s) (pm)	E2 NBO energies (kJ/mol)	NH(i) stretch frequency (cm^{-1})
One-residue compounds						
Cbz-Cys(Me)-NHMe **1**	5-6$^{\gamma}$ **1-a (g-)**	1 2	5 6$^{\gamma}$	216 245	7.2 30.5	3438 3346
Cbz-Attc-NHMe [a]	5-6$^{\gamma}$ **1 (t)**	1 2	5 6$^{\gamma}$	204 240	13.9 32.3	3400 3327
Two-residue compounds						
Cbz-[Cys(Me)]$_2$-NHMe **2** A	5$^{\gamma}$ - 5$^{\gamma}$/π_{am1} - 10/π_{am2} **1-b (*g+*)**	1 2 3	5$^{\gamma}$ 5$^{\gamma}$ / π_{am1} 10 / π_{am2}	262 259 / 230 219 / 231	8.6 11.7 / 2.8 14.1 / 2.2	3420/3417[b] 3404 3420/3417[b]
Cbz-[Cys(Me)]$_2$-NHMe **2** B	5$^{\gamma}$ - 5$^{\gamma}$/π_{am1} - 10/π_{am2} **1-c (*g+*)**	1 2 3	5$^{\gamma}$ 5$^{\gamma}$ / π_{am1} 10 / π_{am2}	258 265 / 232 208 / 234	21.8 9.5 / 3.3 8.2 / 1.6	3392 3410 3417
Cbz-[Cys(Me)]$_2$-NHMe **2** C	5$^{\gamma}$ - 5$^{\gamma}$/π_{am1} - 7 **2-c (*g+*)**	1 2 3	5$^{\gamma}$ 5$^{\gamma}$/π_{am1} 7	275 257 / 226 195	2.8 8.9 / 3.4 36.8	3443 3413 3322
Cbz-(Attc)$_2$-NHMe A [a]	5-6$^{\gamma}$/5-6$^{\gamma}$ **1 (t)**	1 2 3	5 6$^{\gamma}$/5 6$^{\gamma}$	204 248 / 206 243	13.0 26.9 /19.1 29.2	3414 3231 3349
Cbz-(Attc)$_2$-NHMe B [a]	f-7$_D$-7$_D$ **4 (g-)**	1 2 3	free 7 7	 198 200	- 25.6 25.1	3486 3368 3332

a : ref (Imani et al. 2020)
b : coupled modes : 5$^{\gamma}$ and 10

Gas phase characterization

The near UV gas phase spectrum of the phenyl chromophore of **1**, depicted in Fig.5, exhibited a few narrow features (bandwidth of the order of a few cm^{-1}, limited by the rotational band contour at ~10 K), which illustrated the rotational and vibrational cooling achieved in the expansion. The IR/UV spectra recorded in the amide A region from the two intense UV bands A and B were similar but sufficiently different (Fig. 5 and Fig. S5) for us to assign them as originating from the electronic ππ* transition of the phenyl group in two distinct conformers, labelled A and B. The A_1 and B_1 bands were found to yield amide A IR/UV spectra superimposable with those obtained from A and B respectively, and were thus assigned to vibronic bands associated to a Franck-Condon activity involving a low frequency (18 cm^{-1}) torsional mode of the Cbz cap (as previously observed in the case of the closely related Attc residue (Imani et al. 2020)). The IR spectra of A and B (Fig. 5 and S5) were composed of a well resolved doublet in each of the regions recorded, amide I, II and A; the latter, observed in the 3400 cm^{-1} region, corresponding to medium range H-bonds (Brenner et al. 2019; Gloaguen et al. 2020). The assignment of conformers A and B was carried out by comparing the experimental data with the IR theoretical spectra of each backbone family, in the three amide A, I and II regions recorded (Fig. S6 and Table S1). This comparison showed that backbone families 1-a, 1-c and 1-d (C5-C6γ backbone) achieved a fair fit both in terms of frequency and intensity pattern, especially in the amide I and II stretch regions which are most sensitive to the backbone conformation, due to the structure-dependent coupling of the carbonyl groups. A more precise assignment on this basis alone was difficult due to significant deviations in the amide A region, which is most sensitive to the H-bonding, including the NH⋯S interactions. However due to the sharp difference in terms of stability between families 1-a, 1-c and 1-d, conformers A and B were assigned to the most stable backbone family 1-a, with a [*t*,*g*+] side chain. The significant disagreement between the observed and theoretical frequencies in the amide A, a feature already noticed in the case of the rigid related compound Attc (see below), was assigned to a weakness of the level of theory used to properly account for the C6γ NH⋯S H-bonds. Regarding Cbz-rotamerism, Fig. S5 (and Table S1) illustrates the similarities in the theoretical spectra of the three Cbz rotamers of this 1-a family, which did not enable us to provide a more precise assignment of A and B.

A point-by-point comparison of these spectroscopic features for **1** and for Cbz-Attc-NHMe is provided in Figure 5. The clear similarity between the gas phase UV or IR spectra illustrates the shared C5-C6γ backbone, characterized by a strong C6γ H-bond (see E2 NBO energies in Table 1). The more flexible side chain of **1** leads to slightly weaker H-bonds (as indicated by the smaller red-shift of the NH⋯S band in **1**), which underlines the strong constraints exerted by the stiff thietane ring of Attc, in particular on the C6γ H-bond. As expected, the theoretical H-bonding distances obtained (Fig. 5) vary accordingly between both compounds.

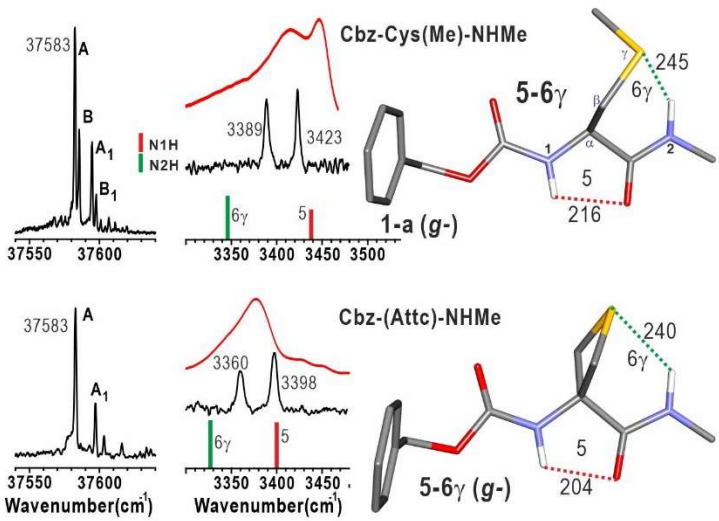

Fig. 5. Top panel. (*left*) Gas phase UV spectra of jet-cooled compound **1**. (*centre*) Gas phase IR spectrum (black) obtained on UV band A, and solution phase spectrum (red) of compound **1**, compared to the theoretical IR spectrum (stick) of the Cbz *g*– rotamer of the lowest energy conformation. (*right*) Corresponding calculated lowest energy conformation. **Bottom panel.** The same spectral data for the corresponding Attc derivative, with the structure of the *g*- Cbz rotamer of the lowest energy conformation.

Solution characterization

The amide A IR absorption spectrum of a $CHCl_3$ solution of compound **1** displayed two amide NH stretch bands, centered at 3418 cm^{-1} and 3445 cm^{-1} (Fig.5). The broad band at 3418 cm^{-1} was attributed largely to the C5-C6γ conformer (family 1, Fig. 4). The band incorporates overlapping absorptions from NH(1) in a C5 intra-residue interaction and NH(2) in a C6γ NH···S interaction. The analogous C5-C6γ conformer was the only significant structure previously observed for Cbz-Attc-NHMe in solution (Fig. 5). The frequencies are blue-shifted for **1** compared to those for Cbz-Attc-NHMe as a result of the longer H-bonding distances. The sharper band at 3445 cm^{-1} in the spectrum of **1** was attributed to the free NH(2) vibration in a C5γ-π_{am1} conformer (family 2, Fig. 4), although this conformer was not observed in the gas phase. It's NH(1) is involved in a C5γ intra-residue NH···S interaction and contributes to the broad band centered at 3418 cm^{-1}. A contribution from a third conformer of **1** with a C5γ-C7 structure (family 3, Fig. 4) cannot be ruled out, as the C5γ interaction of NH(1) would be expected to contribute to the broad band at 3418 cm^{-1} while the C7 interaction of NH(2) might explain the slight shoulder observed at around 3345 cm^{-1}.

The absorption bands of **1** in the amide I and II region (Fig. S7) showed good general correlation with the gas phase data with a slight red-shift (*c* 20 cm^{-1}) for the amide I (C=O stretching) bands. These latter were observed at 1719 cm^{-1} and 1679 cm^{-1} and correspond to a free carbamate and a C5 bonded amide, respectively. These data are in excellent agreement with those reported previously for $CHCl_3$ solutions of Cbz-Aib-Xaa-OMe (Xaa = Aib or Ala) in which the N-terminal Aib adopted a C5 feature leaving the carbamate C=O free (bands at 1724 cm^{-1} and 1675 cm^{-1}, respectively).(Rao et al. 1980) As expected, the data for Cbz-Attc-NHMe in this region are very similar to those observed for **1** (Fig. S7).

Compound **1** was also examined in solution in $CDCl_3$ (5 mM) by ^1H NMR spectroscopy, which provides an indication of the average behavior of each NH (Fig. S1). The principal signals for both the carbamate NH(1) (δ 5.82 ppm) and the amide NH(2) (δ 6.57 ppm) were accompanied by minor satellite signals, consistent with the presence of more than one conformer (or rotamer). The signals were moderately deshielded, suggesting that H-bonding was less strong and/or prevalent than in Cbz-Attc-NHMe, for which the corresponding signals appeared at 6.45 ppm and 7.98 ppm, respectively. Titration of the solution of **1** with DMSO-d_6 (Section S3.1.) induced a moderate downfield shift for each signal (Δδ 0.52 ppm and 0.74 ppm, respectively, for 10% added DMSO). These values pointed to partial solvent exposure but indicated that the conformational landscape implicated H-bonding, at least to some extent, of both NHs. These data are in agreement with the proposal of the coexistence of two (and possibly three) conformers, C5-C6γ and C5γ-π_{am1} (and possibly C5γ-C7), for **1** in solution.

Cbz-[Cys(Me)]$_2$-NHMe 2

Quantum chemistry investigation

The conformational landscape of **2** at 300 K (red levels in Fig. 6) revealed that the most stable backbone family, labelled 1, corresponded to type I β-turns, featuring a C10 H-bond bridging the NH(3) group with the carbamate CO, and two intra-residue C5γ H-bonds bridging the NH(1) and NH(2) groups to the side chain S atom of the same residue. The subfamilies 1-a to 1-j differed by the side chains' orientations, which gave rise to significant differences in C5γ H-bond strengths (Table S2). In these β-turn structures, the C10 bond remained modest, as shown by the NBO indicators (~10 kJ/mol ; Table 1), with elongated H-bonding distances (210-220 pm range), in line with their frustrated character due to the stiffness of the turn backbone (Brenner et al. 2019). Interestingly, the compactness of the turn structure also enabled weak ancillary π_{amide} interactions (~3 kJ/mol), which contributed marginally to the stability of the architecture. The turn structure is geometrically amenable to the formation of intra-residue C5γ bonds. The flexibility of the side chains, however, enabled several orientations, which were influenced by minor interactions, such as S-Me···Me-S or S-Me···phenyl interactions. As a result the C5γ interactions were quite diverse in strength (8-22 kJ/mol range,

Table 1). They were much weaker and more elongated than the C6γ bonds observed in conformers of compound **1** (NBO stabilization energy of ca. 30 kJ/mol; Table 1), which are forbidden here due to their incompatibility with a C10 folded backbone. As for the C10 bond, the relatively large distance and modest strength, compared to unconstrained H-bonds, such as those in an intermolecular complex (Goldsztejn et al. 2020b), clearly illustrated the frustration of these H-bonds within family 1, constrained by the stiffness and the shortness of the Cys(Me) side chain.

The other families (2-5) corresponded to sequential combinations of structures observed in compound **1**, exhibiting either C5, C7 or π_{amide} H-bonds together with C5γ or C6γ NH···S bonds. They were higher in energy, by at least 10 kJ/mol, precluding their observation as major conformations at room temperature. This landscape suggests that family 1 should be the most populated in the gas phase at high temperature. Regarding rotamer populations, the enhanced stability of the rotamers *g*+ observed at 0 K (Fig. 6), due to the occurrence, in this phenyl ring disposition, of dispersive interactions favoring a close contact between the side chain methyl and the phenyl ring, suggested that conformational relaxation should funnel the population to those rotamers.

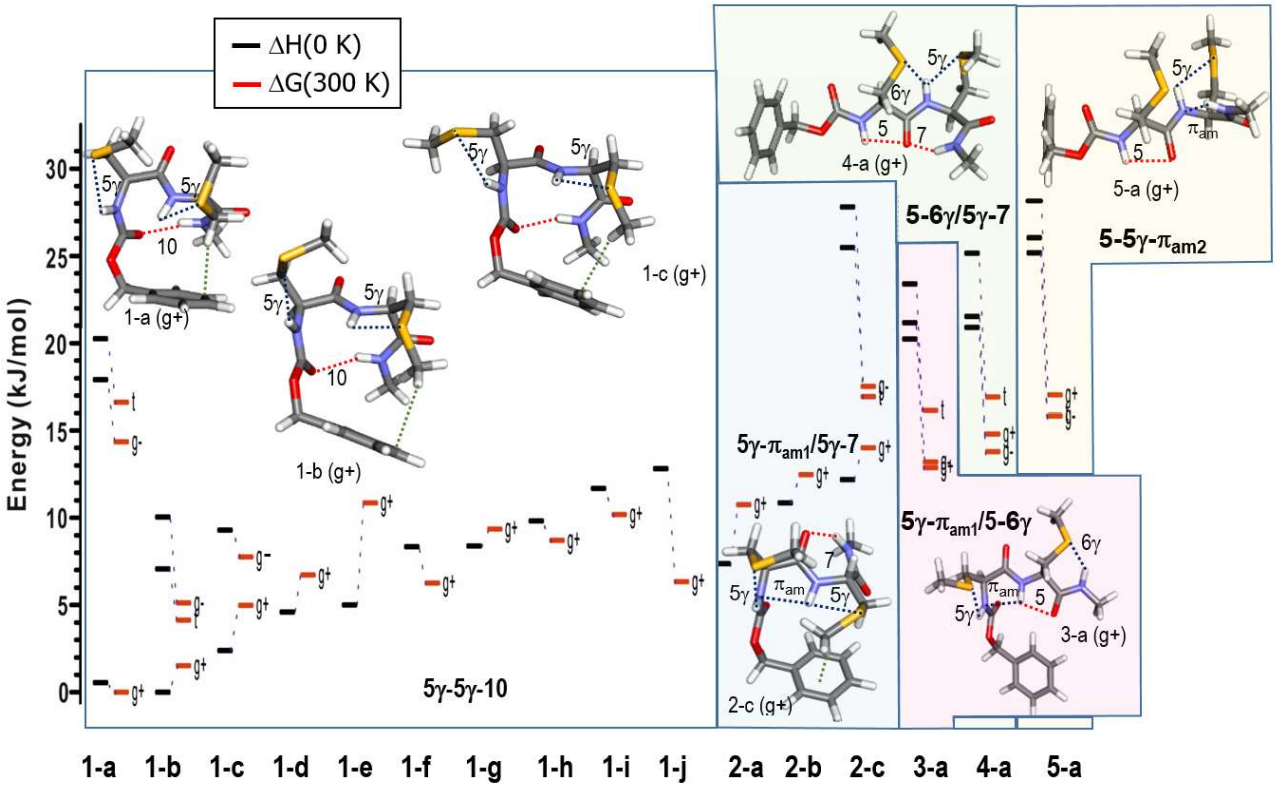

Fig. 6. Conformational landscape of **2** at RI-B97-D3-BJ(abc)/def2-TZVPPD level of theory, organized according to backbone families. Within each family, side chain orientations can be found (see SI for structure details). For (sub)families 1 and 2 of high energy at 300 K, only the *g*+ Cbz rotamer having the Cbz side chain contact is shown. For the sake of clarity, the π_{amide} interactions of the β-turns have been omitted.

Gas phase characterization

The near UV gas phase spectrum of the phenyl chromophore of **2**, depicted in Fig.7, again exhibited two intense narrow features. IR/UV spectroscopy (Fig. 7, top) enabled us to assign them to different conformers, labelled A and B. Several much weaker UV bands were also observed: two of them, A_1 and B_1, were assigned to the same conformer as A and B respectively (again due to identical amide A IR spectra); a third band (C) was assigned to another conformer labelled C (Fig. 7). The remaining still-weaker bands, accounting for less than 15 % of the total population, were too weak to be probed using the present technique. The IR spectra were composed of a well-resolved triplet in each of the regions recorded, amide A (Fig. 7) and amide I, II (Fig. S8 and S9). The IR/UV spectra of A and B were found to be similar in the amide I and II regions but differed

significantly in the NH stretch range (Fig. 7), suggesting similar backbones but different side chain H-bonding patterns; the 3400 cm^{-1} region pattern showed evidence for relatively weak H-bonds, compared for example to C7 H-bonds (Gloaguen and Mons 2015; Imani et al. 2020).

Theoretical spectra (Figs. 7, S8 and S9) again illustrated the efficient diagnostics of the backbone structure provided by the amide I and II regions. Comparison with experimental spectra demonstrated that conformers A and B adopt a type I β-turn structure, in contrast to C. A detailed examination of the triplet features (frequency and intensity pattern) of the most stable rotamers in each backbone sub-family (Table S2 and Fig. S8) indeed showed that the *g*+ rotamers of sub-families 1-b and 1-c best fit to conformers A and B respectively (Fig. 7). The absence of a representative of sub-family 1-a, despite its stability, was then ascribed to significant conformational relaxation within the β-turn family basin (1) in the late expansion. The difference in pattern between the NH stretch triplets of A and B was ascribed to a different orientation of the first side chain (Tables 1 and S2), which caused the NH···S bands to change accordingly to the H-bonding distances, whereas little variation was noted for the β-turn C10 H-bond, found at 3410 cm^{-1} and 3415 cm^{-1}, respectively.

Regarding conformer C, the splitting of the amide I triplet (Fig. S9) pleaded in favor of the presence of a C7 H-bond. A careful examination of the theoretical spectra within family 2 provided several good candidates, with low energy conformations, leading us to assign this conformer to a C5γ-π_{am1}/C5γ-C7 backbone, the best fit being achieved for the 2-c (g+) conformer, although this latter was not the most stable of its family. It should be noted that such a backbone is topologically close to the C5γ-C5γ-C10 backbone of family 1, suggesting that the absence of the most stable structures of family 2 could arise from 2→1 interfamily isomerization during the early expansion; some family 2 structures eventually funneling into family 1 manifold.

The conformational gas phase preferences of **2** and of Cbz-$(Attc)_2$-NHMe (Fig. 7) are remarkably different, in sharp contrast with the close comparison between **1** and Cbz-Attc-NHMe noted above. These differences in backbone preferences (folded C5γ-C5γ-C10 for **2** vs. extended C5-C6γ/C5-C6γ for Cbz-$(Attc)_2$-NHMe) illustrate the emergence of strikingly different backbone structures in the dimers, accompanied by a radical switch between the two types of NH···S H-bonds facilitated by the flexible Cys(Me) side chain.

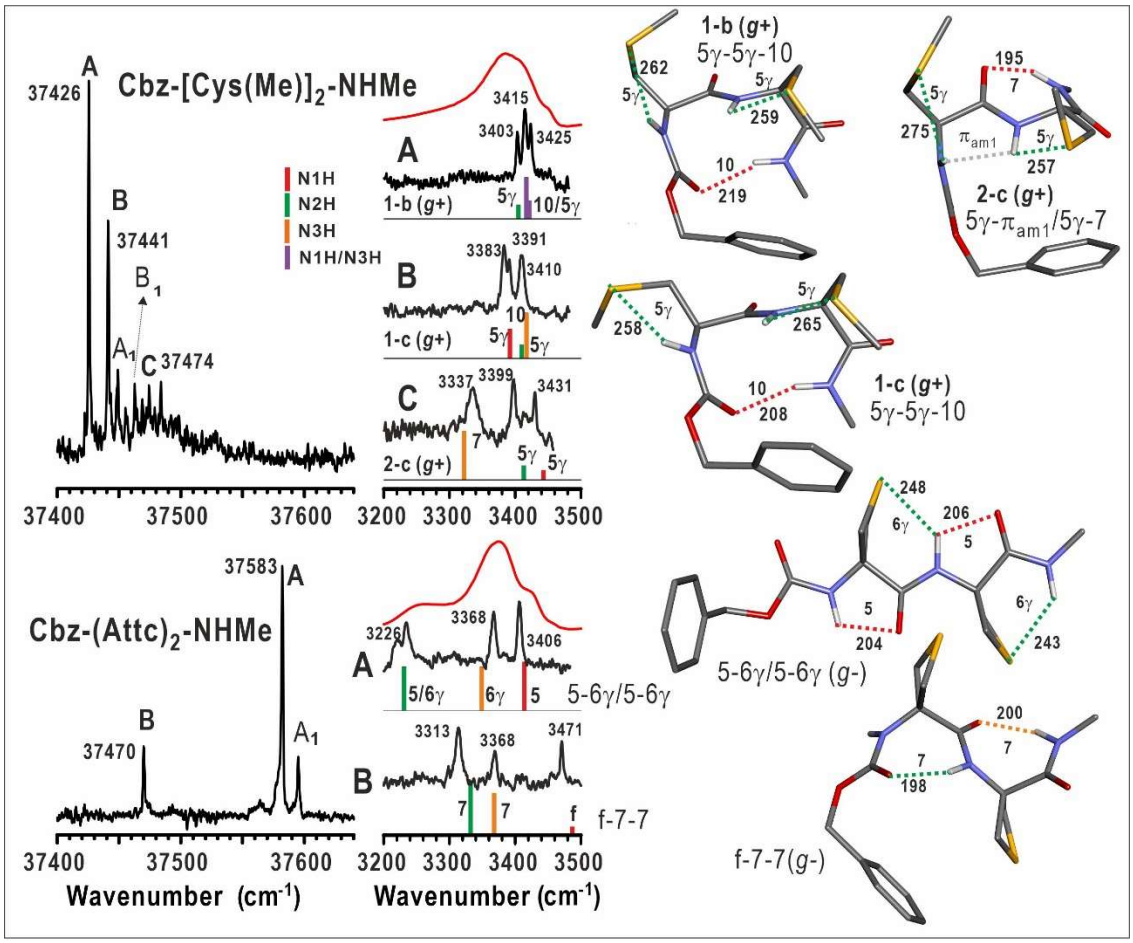

Figure 7 : **Top panel** (*left*) Gas phase UV spectra of jet-cooled compound **2**. (*centre*) Gas phase IR spectrum (black) obtained on UV bands A, B and C and solution phase spectrum (red) of compound **2**, compared to the theoretical IR spectrum (stick) of the Cbz *gauche+* rotamer of the lowest energy conformations of sub-family 1b, 1c and 2c (see Fig. 6). For the sake of clarity, the π_{amide} interactions of the β-turns have been omitted. (*right*) Corresponding calculated conformations. **Bottom panel** Corresponding features for the related Attc compound.

Solution characterization

The amide A region of the IR absorption spectrum of a $CHCl_3$ solution of compound **2** showed one broad NH stretch band, with a maximum at 3387 cm^{-1} with a shoulder at 3409 cm^{-1} (Fig. 7), which correlated well with the gas phase data and was fully consistent with the dominant presence of folded C5γ-C5γ-C10 conformers (family 1, Fig. 7). A minor contribution from a C5γ-π_{am1}/C5γ-C7 conformer (family 2, Fig. 7) cannot be ruled out and may explain the small shoulder around 3449 cm^{-1} and the tailing-off of the main absorption band towards lower frequencies.

In the amide I and II region, IR absorption bands of **2** (Fig. S10) were generally consistent with the gas phase data for the C5γ-C5γ-C10 conformer family, with a slight red-shift (*c* 20 cm^{-1}) for the two amide I bands, observed at 1719 cm^{-1} and 1673 cm^{-1}. The frequency of the former band is in excellent agreement with the C10-bonded carbamate C=O absorption (1718 cm^{-1}) for Cbz-$(Aib)_3$-OMe in $CHCl_3$ solution (Pulla Rao et al. 1980). Nonetheless, the close similarity of the spectral data for **2** and for Cbz-$(Attc)_2$-NHMe in the amide I and II region, despite the difference in the conformational preferences of these two compounds, indicates that reliable differentiation between C5 and C10 conformers is not as straightforward here as in other cases.(Peggion et al. 2013)

Examination of **2** by ^{1}H NMR spectroscopy in $CDCl_3$ (5 mM) provided good support for the β-turn folded C5γ-C5γ-C10 conformer (Fig. S3). All three NHs, carbamate NH(1) (δ 5.80 ppm), amide NH(2) (δ 7.14 ppm) and amide NH(3) (δ 6.91 ppm), were partially deshielded, while titration with DMSO-d_6 induced moderate downfield shifts for each signal (Δδ 0.66 ppm, 0.65 ppm, and 0.36 ppm, respectively, for 10% added DMSO; see details in Section S3.2). These data are compatible with the presence of a C5γ-C5γ-C10 conformer (family 1 described above), whose C10 H-bonded NH(3) leads to a low DMSO-d_6 titration coefficient.

Further evidence was obtained from a NOESY experiment, which revealed long-range correlations between the C-terminal amide CH_3 protons and the aryl protons (and to a lesser extent the benzylic protons) of the Cbz group (Fig. 8). These interactions can only arise from a β-turn conformer with a C10 feature. Similar correlations were observed between the side chain methyl groups of each Cys(Me) residue and the aryl protons and one of the benzylic protons of the Cbz group, which is consistent with a globally folded conformer which can accommodate NH···S interactions. Indeed, a correlation was observed between NH(1) and the methyl group of the first Cys(Me) residue (not illustrated) although an analogous correlation for the second residue was not observed. Collectively these correlations support the presence of a C10 conformer which may also adopt C5γ interactions. These data contrast significantly with the NOESY data obtained previously for Cbz-(Attc)$_2$-NHMe, which showed only short-range correlations consistent with a fully extended conformation (Imani et al 2020).

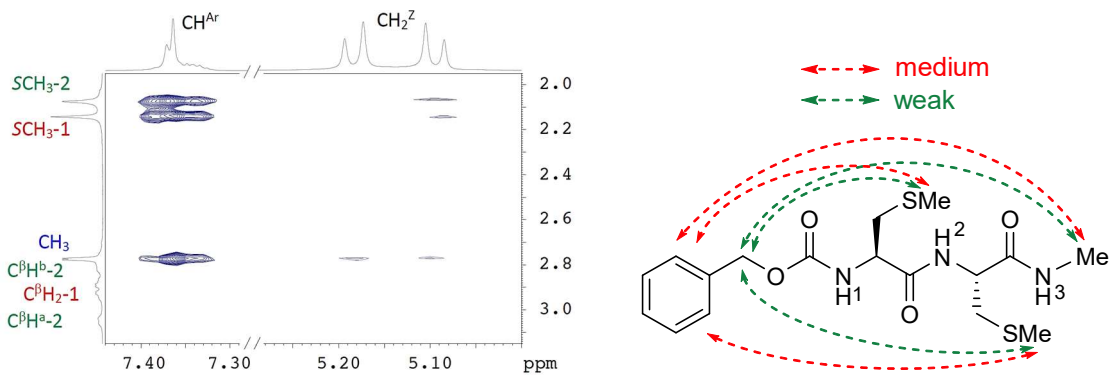

Fig. 8. Significant long-distance NOESY correlations observed for compound **2**. Shorter distance correlations have been omitted for clarity.

Conclusions

In the light of the above results, the nature of NH···S H-bonding in small oligomers of Cys(Me) can be rationalized. It had previously been observed that in polar solvents such compounds were largely disordered (Toniolo et al. 1975). However, in gas phase and in low polarity solution, the model compound for one Cys(Me) residue, **1**, adopts a preferred C5-C6γ conformation which combines an intra-residue N−H···O=C backbone interaction (C5) and an inter-residue N−H···S interaction implicating the side-chain sulfur atom (C6γ). This structural feature is similar to that exhibited by the thietane amino acid Attc, in its model compound Cbz-Attc-NHMe, although the H-bonds formed by Cys(Me) are noticeably weaker due to the absence of strong constraints on the acyclic side chain. In addition, other low energy structures incorporating an alternative N−H···S interaction, namely the intra-residue C5γ, may contribute to the conformational landscape of **1**, in contrast with Attc, where the C5γ interactions cannot form due to the strong covalent constrains of the thietane ring, which penalises these structures by *c* 5 kJ/mol. Interestingly, in the closely related compound Cbz-Cys-NH_2, backbone structures with either C5γ or C6γ N−H···S interactions were reported recently.(Goldsztejn et al. 2020a)

In the Cys(Me) dimer model, **2**, the dominant conformation is a regular type I β-turn. This structure is incompatible with concomitant C6γ interactions and in consequence they are no longer in evidence. Instead,

C5γ interactions which are fully compatible with the turn geometry, occur and additionally stabilize the structure. This behavior is in marked contrast with the corresponding model compound for Attc, Cbz-$(Attc)_2$-NHMe which adopts an extended C5-C6γ/C5-C6γ conformation. The theoretical studies predicted that such a conformer for **2** would be around 20 kJ/mol higher in energy than the β-turn (despite the modest H-bond strengths in the latter), presumably because of the missing covalent constraints provided by the thietane ring of Attc. The switch in behavior of the Cys(Me) dimer therefore signals the onset of the coexistence of N–H⋯S interactions (C5γ) with other backbone amide interactions (C10 in the present case) which are compatible with regular secondary structures, as was previously postulated for longer oligomers of this amino acid.

The present work emphasizes the role of the side chain flexibility onto the conformational preferences of oligomers of amino acids exhibiting an NH⋯S H-bonding. This matter will be investigated further by our consortium, using the full panel of techniques presently showcased, by varying the size of the sulfur-containing ring of cyclic amino acids.

Acknowledgements

The authors wish to thank Ms Roxanne Berthin and Ms Anna Kriukova for their theoretical contribution to the conformational explorations. Support from the French National Research Agency (ANR; Grant ANR-17-CE29-0008 "TUNIFOLD-S") and from the “Investissements d’Avenir” Funding program (LabEx PALM; grant ANR-10-LABX-0039-PALM; DIRCOS) are acknowledged. This work was granted access to the HPC facility of [TGCC/CINES/IDRIS] under the Grant 2019-A0050807540 awarded by GENCI (Grand Equipement National de Calcul Intensif) and to the CCRT High Performance Computing (HPC) facility at CEA under the Grant CCRT2019-p606bren.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants and/or animals This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent This article does not contain individual participants requiring informed consent.

References

Alabugin IV, Gilmore KM, Peterson PW (2011) Hyperconjugation. WIREs Comput Mol Sci 1 (1):109-141. doi:10.1002/wcms.6

Alabugin IV, Gomes GD, Abdo MA (2019) Hyperconjugation. WIREs Comput Mol Sci 9 (2). doi:10.1002/wcms.1389

Alauddin M, Biswal HS, Gloaguen E, Mons M (2015) Intra-residue Interactions in Proteins: Interplay Between Serine or Cysteine Side Chains and Backbone Conformations, Revealed by Laser Spectroscopy of Isolated Model Peptides. Physical Chemistry Chemical Physics 17 (3):2169-2178. doi:10.1039/c4cp04449e

Amin MA, Khaled KF, Mohsen Q, Arida HA (2010) A study of the inhibition of iron corrosion in HCl solutions by some amino acids. Corrosion Sci 52 (5):1684-1695. doi:10.1016/j.corsci.2010.01.019

Birdi KS, Fasman GD (1972) Monolayer Studies of Synthetic Poly(Alpha-Amino Acids). Journal of Polymer Science Part a-1-Polymer Chemistry 10 (8):2483-&. doi:10.1002/pol.1972.150100824

Biswal HS, Gloaguen E, Loquais Y, Tardivel B, Mons M (2012) Strength of NH···S Hydrogen Bonds in Methionine Residues Revealed by Gas-Phase IR/UV Spectroscopy. Journal of Physical Chemistry Letters 3 (6):755-759. doi:10.1021/jz300207k

Bloom SM, Deloze C, Fasman GD, Blout ER (1962) Effect of Amino Acid Composition on Conformations of Synthetic Polypeptides, Polymers and Copolymers of L-Methionine S-Methyl-L-Cysteine and L-Valine. J Am Chem Soc 84 (3):458-&. doi:10.1021/ja00862a027

Bonora GM, Maglione A, Toniolo C (1975) Linear Oligopeptides. 24. Preparation of a Series of Monodisperse Homopeptides (to Heptamer) of S-Methyl-L-Cysteine. Gazz Chim Ital 105 (9-10):1055-1062

Brenner V, Gloaguen E, Mons M (2019) Rationalizing the Diversity of Amide-Amide H-Bonding in Peptides using the Natural Bond Orbital Method. Physical Chemistry Chemical Physics 21 (44):24601-24619. doi:10.1039/c9cp03825f

D. Cavallini, G. E. Gauli, V. Zappia (1980). In: Natural Sulfur Compounds: Novel Biochemical and Structural Aspects. Plenum Press, New York

Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. Theoretical Chemistry Accounts 97 (1-4):119-124. doi:10.1007/s002140050244

Fasman GD, Potter J (1967) Optical Rotatory Dispersion of 2 β Structures. Biochem Biophys Res Commun 27 (2):209-&. doi:10.1016/s0006-291x(67)80063-9

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, R. Gomperts R, B. Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai N, Vreven T, Throssell K, Montgomery Jr. JA, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian 16. Revision A.03 edn. Gaussian, Inc., Wallingford CT.

Gloaguen E, Brenner V, Alauddin M, Tardivel B, Mons M, Zehnacker-Rentien A, Declerck V, Aitken DJ (2014) Direct Spectroscopic Evidence of Hyperconjugation Unveils the Conformational Landscape of Hydrazides. Angewandte Chemie-International Edition 53 (50):13756-13759. doi:10.1002/anie.201407801

Gloaguen E, de Courcy B, Piquemal JP, Pilme J, Parisel O, Pollet R, Biswal HS, Piuzzi F, Tardivel B, Broquier M, Mons M (2010a) Gas-Phase Folding of a Two-Residue Model Peptide Chain: On the Importance of an Interplay between Experiment and Theory. J Am Chem Soc 132 (34):11860-11863. doi:10.1021/ja103996q

Gloaguen E, Mons M (2015) Isolated Neutral Peptides. Topics in Current Chemistry 364:225-270. doi:10.1007/128_2014_580

Gloaguen E, Mons M, Schwing K, Gerhards M (2020) Neutral Peptides in the Gas Phase : Conformation and Aggregation Issues. Chemical Reviews 120 (22):12490–12562. doi:10.1021/acs.chemrev.0c00168

Gloaguen E, Valdes H, Pagliarulo F, Pollet R, Tardivel B, Hobza P, Piuzzi F, Mons M (2010b) Experimental and Theoretical Investigation of the Aromatic-Aromatic Interaction in Isolated Capped Dipeptides. Journal of Physical Chemistry a 114 (9):2973-2982. doi:10.1021/jp904216f

Goldsztejn G, Mundlapati V-R, Donon J, Tardivel B, Gloaguen E, Brenner V, Mons M (2020b) An intraresidue H-bonding motif in selenocysteine and cysteine, revealed by gas phase laser spectroscopy and quantum chemistry calculations. Physical Chemistry Chemical Physics 22 (36):20409-20420. doi:10.1039/D0CP02825H

Goldsztejn G, Mundlapati V, Brenner V, Gloaguen E, Mons M, León I, Cabezas C, Alonso JL (2020a) Intrinsic Folding of the Cysteine Side Chain : Microwave and Optical Spectroscopic Studies Combined with Quantum Chemistry Methods. Physical Chemistry Chemical Physics 22:20284-20294. doi:10.1039/D0CP03136D

Grabarse WG, Mahlert F, Shima S, Thauer RK, Ermler U (2000) Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: Unusual amino acid modification, conservation and adaptation. Journal of Molecular Biology 303 (2):329-344. doi:10.1006/jmbi.2000.4136

Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. Journal of Chemical Physics 132 (15). doi:10.1063/1.3382344

Imani Z, Mundlapati VR, Goldsztejn G, Brenner V, Gloaguen E, Guillot R, Baltaze JP, Le Barbu-Debus K, Robin S, Zehnacker A, Mons M, Aitken DJ (2020) Conformation control through concurrent N-H···S and N-H···O-C hydrogen bonding and hyperconjugation effects. Chemical Science 11 (34):9191-9197. doi:10.1039/d0sc03339a

Komoto T, Oya M, Kawai T (1974) Crystallization of Polypeptides in Course of Polymerization. 6. Effect of Heteroatom Attached to β-Carbon on Crystal-Growth. Makromolekulare Chemie-Macromolecular Chemistry and Physics 175 (1):301-310. doi:10.1002/macp.1974.021750130

Lei M, Feng HY, Bai EH, Zhou H, Wang J, Qin YR, Zhang HY, Wang XY, Liu ZG, Hai O, Liu J, Zhu YQ (2019) Discovery of a novel dipeptidyl boronic acid proteasome inhibitor for the treatment of multiple myeloma and triple- negative breast cancer. Organic & Biomolecular Chemistry 17 (3):683-691. doi:10.1039/c8ob02668h

Liu PX, Weng R, Sheng XJ, Wang XL, Zhang WH, Qian YZ, Qiu J (2020) Profiling of organosulfur compounds and amino acids in garlic from different regions of China. Food Chemistry 305. doi:10.1016/j.foodchem.2019.125499

Macromodel (Schrödinger Release 2019-3). Schrödinger, LLC,, New York, NY

Maw GA (1982) Biochemistry of S-Methyl-L-Cysteine and its Principal Derivatives. Sulfur Reports 2 (1):1-26. doi:10.1080/01961778208082422

Mazumder MAJ (2019) Synthesis, characterization and electrochemical analysis of cysteine modified polymers for corrosion inhibition of mild steel in aqueous 1 M HCl (vol 9, pg 4277, 2019). Rsc Advances 9 (11):6077-6077. doi:10.1039/c9ra90012h

Palumbo M, Darin S, Bonora GM, Toniolo C (1976) Linear Oligopeptides. 29. Infrared Conformational-Analysis of Homo-Oligopeptides in Solid-State and in Solution. Makromolekulare Chemie-Macromolecular Chemistry and Physics 177 (5):1477-1492. doi:10.1002/macp.1976.021770519

Peggion C, Moretto A, Formaggio F, Crisma M, Toniolo C (2013) Multiple, Consecutive, Fully-Extended 2.0(5)-Helix Peptide Conformation. Biopolymers 100 (6):621-636. doi:10.1002/bip.22267

Rao CP, Nagaraj R, Rao CNR, Balaram P (1980) Infrared Studies on the Conformation of Synthetic Alamethicin Fragments and Model Peptides Containing Alpha-Aminoisobutyric-Acid. Biochemistry 19 (3):425-431. doi:10.1021/bi00544a004

Rappoport D, Furche F (2010) Property-optimized Gaussian basis sets for molecular response calculations. Journal of Chemical Physics 133 (13). doi:10.1063/1.3484283

Reed AE, Curtiss LA, Weinhold F (1988) Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint. Chemical Reviews 88 (6):899-926. doi:10.1021/cr00088a005

Thompson JF (1967) Sulfur Metabolism In Plants. Annual Review of Plant Physiology 18:59-+. doi:10.1146/annurev.pp.18.060167.000423

Toniolo C (1980) Intramolecularly Hydrogen-Bonded Peptide Conformations. CRC Critical Reviews in Biochemistry 9 (1):1-44. doi:10.3109/10409238009105471

Toniolo C, Bonora GM, Scatturin A (1975) Linear Oligopeptides. 25. Effect of a Sulfur Atom in γ-Position on Secondary Structures of Homo-Oligopeptides. Gazz Chim Ital 105 (9-10):1063-1071

Turbomole (2017) V7.2. A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, Turbomole GmbH, since 2007, available from http://www.turbomole.com.

Weinhold F (2012) Natural Bond Orbital Analysis: A Critical Overview of Relationships to Alternative Bonding Perspectives. J Comput Chem 33 (30):2363-2379. doi:10.1002/jcc.23060

Weinhold F, Reed AE, Carpenter JE, Glendening ED. NBO version 3.1. edn.,

Wróbel J (1977). In: Low Molecular Weight Sulphur Containing Natural Products. Pergamon Press, Oxford,

Yan B, Jaeqx S, van der Zande WJ, Rijs AM (2014) A Conformation-Selective IR-UV Study of the Dipeptides Ac-Phe-Ser-NH_2 and Ac-Phe-Cys-NH_2: Probing the SH···O and OH···O Hydrogen Bond Interactions. Physical Chemistry Chemical Physics 16 (22):10770-10778. doi:10.1039/c4cp00810c

Zawistowski J, Kopec A, Jedrszczyk E, Francik R, Bystrowska B (2018) Garlic Grown from Air Bulbils and Its Potential Health Benefits. In: Jayaprakasha GK, Patil BS, Gattuso G (eds) Advances in Plant Phenolics: From Chemistry to Human Health, vol 1286. ACS Symposium Series. pp 315-328