An intraresidue H-bonding motif in selenocysteine and cysteine, revealed by gas phase laser spectroscopy and quantum chemistry calculations

Gildas Goldsztejn, Venkateswara Rao Mundlapati, Jérémy Donon, Benjamin Tardivel, Eric Gloaguen, Valérie Brenner, Michel Mons

▶ **To cite this version:**

Gildas Goldsztejn, Venkateswara Rao Mundlapati, Jérémy Donon, Benjamin Tardivel, Eric Gloaguen, et al.. An intraresidue H-bonding motif in selenocysteine and cysteine, revealed by gas phase laser spectroscopy and quantum chemistry calculations. Physical Chemistry Chemical Physics, 2020, 22 (36), pp.20409-20420. 10.1039/D0CP02825H . hal-03443065

HAL Id: hal-03443065
https://hal.science/hal-03443065

Submitted on 15 Feb 2024

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.

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2020, **22**, 20409

An intraresidue H-bonding motif in selenocysteine and cysteine, revealed by gas phase laser spectroscopy and quantum chemistry calculations†

Gildas Goldsztejn, ‡ Venkateswara Rao Mundlapati, Jérémy Donon, Benjamin Tardivel, Eric Gloaguen, Valérie Brenner and Michel Mons *

Models of protein chains containing a seleno-cysteine (Sec) residue have been investigated by gas phase laser spectroscopy in order to document the effect of the H-bonding properties of the SeH group in the folding of the Sec side chain, by comparison with recent data on Ser- and Cys-containing sequences. Experimental data, complemented by quantum chemistry calculations and natural bonding orbital (NBO) analyses, are interpreted in terms of the formation of a so-called 5^{γ} intra-residue motif, which bridges the acceptor chalcogen atom of the side chain to the NH bond of the same residue. This local structure, in which the O/S/Se atom is close to the plane of the N-terminal side amide, is constrained by local backbone-side chain hyperconjugation effects involving the S and Se atoms. Theoretical investigations of the Cys/Sec side chain show that (i) this 5^{γ} motif is an intrinsic feature of these residues, (ii) the corresponding H-bond is strongly non-linear and intrinsically weak, (iii) but enhanced by γ- and β-turn secondary structures, which promote a more favorable 5^{γ} H-bonding approach and distance. The resulting H-bonds are slightly stronger in selenocysteine than in cysteine, but nearly inexistent in serine, whose side chain in contrast behaves as a H-bonding donor. The modest spectral shifts of the Cys/Sec NH stretches measured experimentally reflect the moderate strength of the 5^{γ} H-bonding, in agreement with the correlation obtained with a NBO-based H-bond strength indicator. The evolution along the Ser, Cys and Sec series emphasizes the compromise between the several factors that control the H-bonding in a hyperconjugation-constrained geometry, among them the chalcogen van der Waals and covalent radii. It also illustrates the 5^{γ} H-bond enhancements with the Sec and Cys residues favoured by the constraints imposed by the γ- and β-turn structures of the peptide chain.

Received 25th May 2020,
Accepted 3rd August 2020

DOI: 10.1039/d0cp02825h

rsc.li/pccp

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

1 Introduction

Gas phase laser spectroscopic investigations of biosystems and in particular of models of protein chains[1–3] provide interesting insights into the local structure of biomolecules, especially in regions depleted in solvent molecules, such as the inner core of proteins. Recently, the local binding of a side chain to the protein backbone through H-bonding has been investigated for several polar residues by coupling gas phase laser spectroscopy of model molecules and quantum chemistry,[4–9] and compared to conditions in real biosystems such as crystallized proteins[6,7] or fibrils.[10] The gas phase studies of the conformational landscape of Serine (Ser) and Cysteine (Cys) residues embedded in model molecules[4,5] reported a local folding of the side chain controlled by a competition between two types of intraresidue H-bonding interactions labelled 5^{γ} and $^{\gamma}6$ (according to a notation previously introduced for histidine and asparagine polar side chains[6,7]), namely the NH···Y and YH···OC bonds respectively, where Y designates the chalcogen atom of the residue considered (Fig. 1). The structures determined together with the spectral shifts measured (in the -70 cm^{-1} shift range) suggest that the main chain to side chain NH···Y H-bonds are weaker than typical amide–amide intrabackbone interactions, like the 7 H-bonds of γ-turns, or comparable intermolecular H-bonds, such as in the complexes of indole with dimethylether[11] or dimethylsulfide.[12] In the context of a protein chain, the backbone environment is expected to noticeably influence the intramolecular H-bonding, *e.g.*, by hampering an optimal approach of the donor to the acceptor. This local folding, intrinsic to the Ser and Cys side chains, is not strong enough to distort the local backbone

Laboratoire Interactions Dynamiques et Lasers (LIDYL), Université Paris-Saclay, Paris, France. E-mail: michel.mons@cea.fr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cp02825h

‡ Present address: Université Paris-Saclay, CNRS, Institut des Sciences Moléculaires d'Orsay, 91405, Orsay, France.

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

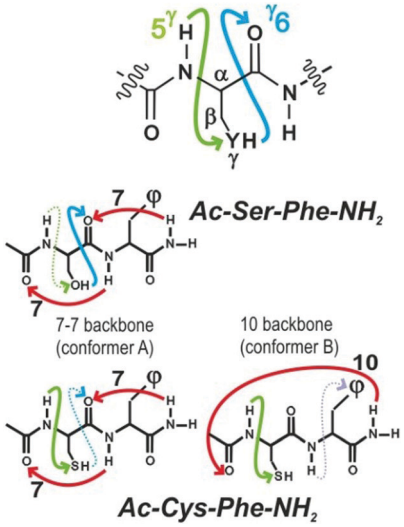

Fig. 1 Top panel: Intraresidue local H-bonding features of the YH group within the series of serine, cysteine and selenocysteine aminoacids in a protein chain. The YH group heteroatom (Y = O, S or Se, respectively) is potentially H-bond acceptor and donor, with corresponding local H-bonds designated with the 5^{γ} and $^{\gamma}6$ symbols respectively. The number indicates the number of atoms in the ring formed by the H-bond and the superscript codes for the position of the heteroatom relative to the Cα atom. Superscript is placed before or after the number, depending on the donor or acceptor character of the side chain respectively. Lower panel: H-bonding schemes observed in the A and B conformers of the Ac-Ser-Phe-NH_2 and Ac-Cys-Phe-NH_2 dipeptides,[5] with 7-7 (double γ-turn) and 10 (β-turn) backbones. Depending on the balance between proton donating and accepting properties of Ser and Cys, $^{\gamma}6$ and 5^{γ} side chain/main chain H-bonds are observed with various relative strengths, as determined from NH stretch spectral shifts (strong and weak H-bonds are figured by full and dotted arrow lines respectively).[5]

structure, which still exhibits the γ- or β-turn secondary structure. However, it can nevertheless influence the conformational distribution by modifying the relative stability of the conformational families.[5]

More precisely, the gas phase comparison[5] of short dipeptide chains containing Ser and Cys residues (Fig. 1, lower panel) showed that β-turn secondary structures (10 backbone) were observed with Cys, whereas they were absent with Ser. This was rationalized by different local side chain H-bonding trends: in Ser, the 7-7 backbone is the only type observed and the $^{\gamma}6$ H-bond originating from the hydroxyl group is the major H-bonding interaction controlling locally the structure; the 5^{γ} H-bond being much weaker. In contrast, the situation is much more open with Cys, where both 7-7 and/or 10 backbones are observed. Both present a local side chain/main chain binding motif, that features a 5^{γ} intraresidue H-bond, linking locally a backbone NH to the sulfur atom lone pair. This is accompanied, in the 7-7 backbone case, by a much weaker $^{\gamma}6$ H-bond, presumably due to the weaker donor character of the SH bond. The presence of such a 5^{γ} structural motif in Cys and the balance between two H-bonding regimes (top of Fig. 1) are expected to depend upon the size and the electronic properties of the Y chalcogen atom, in particular the polar character of the YH bond.

At this stage, it thus appeared to us interesting to investigate whether the somewhat unexpected structuring properties of a relatively apolar residue like Cys, illustrated by this 5^{γ} motif, are also shared by the sister residue selenocysteine (Sec), the analog of cysteine in which the sulfur atom is substituted by selenium. Interestingly this amino acid, considered as the 21st proteinogenic amino acid, is co-translationally encoded by a stop codon (UGA), which differs by only one letter to the cysteine codons (UGU and UGC), making easy mutations interchanging Cys and Sec.[13–16] Sec is also the crucial building block of the selenoproteins, which frequently show catalytic activities on redox reactions[17] and protect cells from damages caused by oxidation.[18] Despite its aforementioned functional importance in proteins,[19] the structural effects of the presence of selenium are seldom addressed.

In this respect, gas phase laser spectroscopy, directly comparable to quantum chemistry calculations, offers an unrivalled approach by providing solvent-free atomic scale data on the proteins' intramolecular interactions. Thus, examples of NH bound to divalent sulfur or selenium atoms have been studied in the gas phase using IR/UV double resonance spectroscopy, where an indole molecule also plays the role of UV chromophore needed by this spectroscopy. The H-bond acceptor is a dimethyl-chalcogenide $Y(Me)_2$ molecule: Y = O;[11] X = S,[12,20] X = Se[21] and the H-bonding marker used is the IR frequency of the NH stretch band in the 3 μm region. Alternatively, amide NH spectral shifts have also been calculated in complexes of *N*-methyl-formamide, taken as a model of a peptide bond, with DMO and DMS.[20,22] It should however be noted that whereas these systems can be considered as good models for a methionine side chain, they cannot pretend to properly model the functional –YH group of the Cys and Sec residue especially in the case of intraresidue H-bonds. To the best of our knowledge, no gas phase approach of the corresponding relevant models was reported so far, in particular on realistic models of protein chain. The present study aims at following the properties of the local 5^{γ} H-bonding motif along the serine-type side chain series, when going down the chalcogen series. We hereafter report a comprehensive study on the preferential conformations adopted by the Sec residue in a model system of a protein chain, namely the *N*-acetyl-selenocysteinyl-phenylalanyl-NH_2 (Ac-Sec-Phe-NH_2) capped dipeptide, with a specific focus on (i) the way the selenium heteroatom does affect the strength of the 5^{γ} H-bonds, (ii) the dependence of the 5^{γ} bond strength upon local backbone folding (γ- *vs.* β- turn), iii) its competition with the other intraresidue bond, the $^{\gamma}6$ H-bond and (iv) its comparison with intermolecular bonding in benchmark complexes. For this purpose, conformer-selective gas phase experimental UV and IR data have been interpreted with the help of quantum chemistry, allowing us to identify the backbone conformations adopted by the Sec residue in the model dipeptide studied, and to characterize their 5^{γ} motif, leading *in fine* to a comparison with Ser and Cys. Moreover, this spectroscopic investigation has been completed by a natural bond orbital (NBO) analysis of the H-bonding and the hyperconjugative interactions that structure the 5^{γ} motif revealing their respective role along the Ac-Xxx-Phe-NH_2 series (Xxx = Ser, Cys, Sec).

 This journal is © the Owner Societies 2020

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

2 Methods

The experimental setup has been detailed elsewhere.[23] Briefly, the sample of interest (Ac-Sec-Phe-NH_2) was mixed with graphite, compressed in a pellet and placed in a vacuum chamber below the nozzle of a pulsed valve. The sample was laser desorbed and the evaporated molecules were entrained and cooled down by the pulsed supersonic jet generated by the expansion of a 7 : 3 He–Ne gas mixture. Then the molecular beam entered the source region of a time-of-flight mass spectrometer (MS), where it interacted with spectroscopy lasers. Neutral molecules were ionized through resonant 2-photon ionization (R2PI), where the first UV photon was used to resonantly excite an electronic transition, thus selecting a particular conformer, and the second photon ionized the excited molecules. The corresponding ions were eventually mass-analyzed and detected in the MS. Few tens of nanoseconds before the UV laser, an IR OPO laser scanned in the NH stretch region ($\sim$3200–3600 cm^{-1}) generated a depopulation of the ion signal whenever the wavelength was in resonance with a vibrational transition of the probed conformer, providing conformer-selective IR spectra of cold species.[3,24]

In order to assign the several conformers probed, we used an approach combining an exploration of the conformation landscape using a molecular mechanics force field with high-level quantum chemistry calculations. The exploration has been carried out using the Macromodel software[25] with the OPLS_2005 force field,[26] while quantum chemistry calculations were performed using the Turbomole package,[27] at the B97-D3 level[28–30] using the Becke–Johnson damping and the three-body term options of Turbomole (B97-D3(BJ)-abc),[27] with a def2-TZVPPD basis set, which proved to be a good compromise between optimized structures, energetics and calculation times in a context where dispersion interactions are ubiquitous.[31] The RI approximation and the auxiliary associated basis sets were used.[32,33] The geometry optimizations were performed using the following parameters: gridsize m3, SCF convergence threshold 10^{-8} u.a., gradient norm convergence threshold 10^{-5} u.a. and a step of 0.02 u.a. for the numerical frequencies. The vibrational frequencies are calculated within the harmonic approximation and scaled through linear scaling functions $f^{\text{scaled}} = a + b{\cdot}f^{\text{harmonic}}$. Mode-dependent scaling parameters (a; b) have been obtained at the present level of theory from the fit of vibrational data on a library[3,31] of gas phase peptides studied in our group: (372.8 cm^{-1}; 0.86953) for the NH, (1209.8 cm^{-1}; 0.63115) for the symmetric NH_2, noted NH_2 s, and (1324.1 cm^{-1}; 0.60872) for the antisymmetric NH_2, noted NH_2 a, stretching modes. 300 K Gibbs energies were calculated using the freeh module of the Turbomole package, since this high temperature energetics has been demonstrated to be relevant to describe conformational population within a supersonic expansion.[3,31]

In order to further characterize the 5^{γ} motif, a NBO analysis was carried out to quantify the stabilizing role of electron delocalization through hyperconjugation and H-bonding effects on the (*N*-(2-thiol/selenol-1-methyl-ethyl)acetamide) model system as well as on the capped dipeptides studied. NBO calculations[34–37] were performed on the RI-B97-D3(BJ)-abc/TZVPPD structures using the NBO module[38] of the Gaussian 09 software.[39] The NBO analysis was carried out at both HF/TZVPP and MP2/TZVPP levels according to previous works, which have shown that this latter level of theory is a good compromise between accuracy and computation times.[40–42]

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 energies through the second order perturbation theory:

$$E(2) = q_i \frac{F(i,j)^2}{|\varepsilon_i - \varepsilon_j|} \quad (1)$$

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 chosen at 0.05 kcal mol^{-1}. In order to characterize the H-bonding, the stabilization energies resulting from the electron delocalization between NBOs of the H-bond proton acceptor and donor, *i.e.*, respectively the donating NBO (Lewis-like occupied lone pairs of the side chain Y atom, Y = O, S or Se, labelled n_Y and n'_Y, or of a backbone carbonyl, n_O, n'_O and π_{CO}) NBOs and the accepting NBO (unoccupied σ^*_{NH} and σ^*_{YH}) respectively, were summed up to provide a total stabilization energy due to the H-bond considered, noted ΣE_{HB}. In the same spirit, significant stabilization energies, corresponding to hyperconjugation effects among the covalent bonds in the vicinity of each motif, and resulting from delocalization between donor and acceptor NBOs, were also considered.

For the sake of comparison with the constrained NH$\cdots$Y H-bonds of Ser, Cys and Sec, the intermolecular complexes of the *trans-N*-methylacetamide with MeOH, MeSH and MeSeH have also been considered theoretically as models of fully unconstrained H-bonds. All the optimized structures are provided in the ESI.†

3 Experimental results and assignments

3.1 UV spectroscopy

The near UV spectrum of Ac-Sec-Phe-NH_2 in the region of the phenyl origin band is displayed in Fig. 2 along with the corresponding spectra of Ac-Ser-Phe-NH_2 and Ac-Cys-Phe-NH_2 previously reported by Alauddin *et al.*[5] Variations along the series of UV spectra reflect the sensitivity of the electronic transitions of the phenyl ring to the conformation adopted by the peptide chain. Experience shows that the effect is sensitive enough to distinguish the spectral features of different peptide chain conformers, appearing as isolated band systems under supersonic expansion conditions.[3] The three spectra present great similarities, in particular the Sec compound exhibits the two systems of bands (labelled A and B) already observed in the Cys counterpart molecule.

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

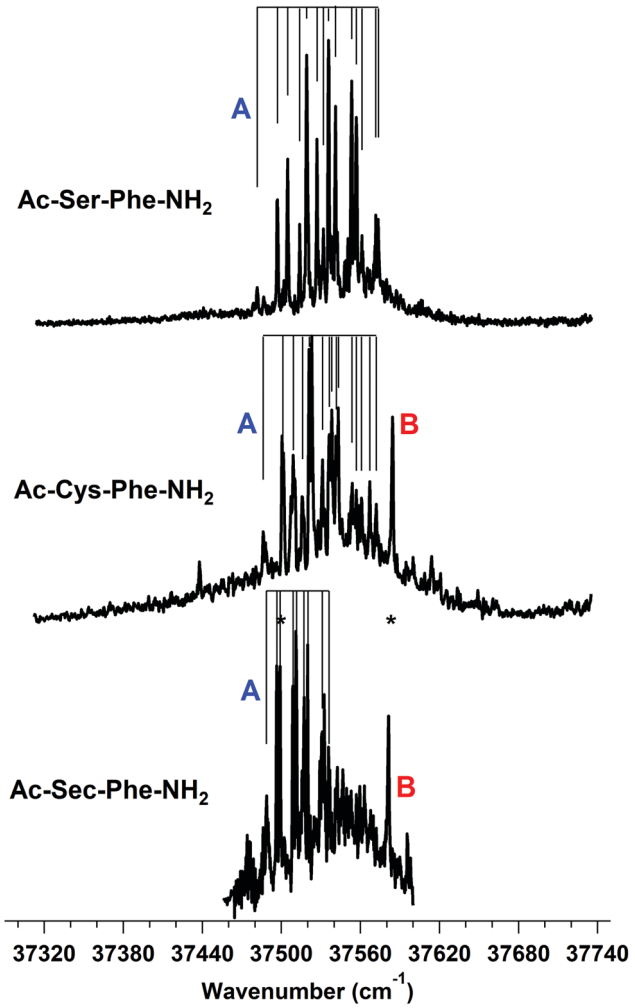

Fig. 2 Mass-selected R2PI spectra of the Ac-Sec-Phe-NH_2 species in the spectral region of the first ππ* transition of phenyl, together with that of the Ser and Cys analogues (adapted from ref. 5), showing the similarity of their UV spectroscopy. An asterisk indicates the bands on which the IR/UV spectra of Fig. 3 were recorded.

The band system A, composed of a rich Franck–Condon vibronic progression, present in both Ser- and Cys-based dipeptides, was previously assigned to a double γ-turn conformer, with a 7-7 backbone H-bond pattern additionally stabilized by $^{\gamma}6$ and/or 5^{γ} local H-bonds on the Ser and Cys residues (Fig. 1).[5] The relative strengths of the 5^{γ} local H-bonds, as estimated from the corresponding NH stretch frequencies measured and the H-bonding distances calculated, were found to depend upon the residue: 5^{γ} being stronger with Cys than with Ser, whereas the $^{\gamma}6$ H-bond exhibited an opposite trend.[5] Alternatively, the isolated band B was assigned to a type I β-turn folded conformer, with a characteristic 10 H-bond, further stabilized by a 5^{γ} local H-bond on the Cys residue (Fig. 1). With Ser, as mentioned in the introduction, the B form is not observed suggesting that this latter type of folding is less favoured.

From the striking similarities between Cys and Sec dipeptides, we can propose a tentative assignment of the A and B spectral features of the Ac-Sec-Phe-NH_2 UV spectrum to the same type of backbone conformations as the Cys counterpart, namely an extended 7-7 and folded 10 backbone.

3.2 IR spectroscopy and structural assignment

First, we measured the IR vibrational spectra in the NH stretch region ($\sim$3200–3600 cm^{-1}) at UV wavelengths shown by vertical lines in Fig. 2 (37 486–37 534 cm^{-1}) and at 37 581 cm^{-1}. Two different IR spectra were observed, confirming that, apart from band B (located at a UV wavelength of 37 581 cm^{-1}), all the intense UV bands probed correspond to a unique conformer A.

These IR spectra were then compared with spectra calculated for the most stable conformations issued from the mixed force field/quantum chemistry exploration. Table 1 shows these conformations, ordered according to the relevant energetics (Gibbs energy) along with the calculated spectra and the corresponding discrepancies with the observed spectra for A and B.

The IR spectrum of conformer A (Top of Fig. 3) is well accounted for by the lowest energy structure that exhibits a double γ-turn backbone (conformation i in Table 1), additionally stabilized by a 5^{γ} H-bond (5^{γ}-7-7 network). The deviations between calculated and experimental frequencies ($\langle\delta\rangle$ = 12 cm^{-1} and δ_{max} = 22 cm^{-1}; see Table 1) are in line with the typical RMS error of the theoretical method ($\sigma \sim 20$ cm^{-1}) that was found empirically on a large number of systems.[3] This double γ-turn structure is characterized by Ramachandran's dihedrals on the Phe residue of $\phi = -83^{\circ}$ and $\psi = 55^{\circ}$ (also see Table S1 of the ESI;† conformer i), and on Cys of $\phi = -88^{\circ}$ and $\psi = 70^{\circ}$, both close to the average inverse γ_L-turn values of $\phi \approx -75^{\circ}$ and $\psi \approx 65^{\circ}$.[43,44] From comparison with the theoretical spectrum (Fig. 3), the red most band is assigned to the stretching mode of the intramolecular C7 $NH_{Phe}\cdots O$ H-bond, the second one to the symmetric stretching mode of the NH_2 group, the third one at 3401 cm^{-1} to the stretching mode of the intramolecular $NH_{Sec}\cdots Se$ bond (noted 5^{γ}) and the last, blue most, band to the antisymmetric stretching mode of the NH_2 group. The IR spectrum of conformer B (Fig. 3 bottom) is well described by the 5^{γ}-π-10 H-bond pattern the most stable structure found in the exploration (conformation a, see Table 1). This structure exhibits Ramachandran dihedrals ($\phi_{Sec} = -70^{\circ}$, $\psi_{Sec} = -9^{\circ}$, $\phi_{Phe} = -94^{\circ}$ and $\psi_{Phe} = 3^{\circ}$; see Table S1 in the ESI†), which fit well those of a type I β-turn ($\phi_{i+1} \approx -60^{\circ}$, $\psi_{i+1} \approx -30^{\circ}$, $\phi_{i+2} \approx -90^{\circ}$ and $\psi_{i+2} \approx 0^{\circ}$).[43] The red most band is assigned to the symmetric stretching mode of NH_2 at 3389 cm^{-1}, which is characteristic of a C10 H-bond;[5] the next one at 3422 cm^{-1} to the stretching mode of the 5^{γ} H-bond between NH and Se and the third one at 3441 cm^{-1} to the stretching mode of NH_{Phe} which is characteristic of a $NH\cdots\pi$ bond with the phenyl ring.[45,46]

3.3 Conformational populations and Cys/Sec comparison

The assignment for the Sec-based compound turns out to be the same as that obtained previously on the Cys-based compound (see comparisons in Table 1). The precision achieved by the theoretical spectra, similar in both cases allows us to consider these assignments with confidence. However, two points deserve discussion regarding the conformational

 This journal is © the Owner Societies 2020

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

Table 1 Theoretical NH stretch frequencies for the most stable calculated conformations of Ac-Sec-Phe-NH_2 within the $\Delta G \leq 8$ kJ mol^{-1} energy range, compared to corresponding experimental features from IR/UV spectra. The calculated conformations are sorted according to their relative Gibbs free energy at 300 K (ΔG); their relative enthalpy at 0 K (ΔH) is also given, as well as the H-bond pattern. For each theoretical conformation, two indices are calculated in order to provide an objective assignment criterion, namely the root-mean square average error $\langle\delta\rangle$ for the four IR bands of the spectrum and the maximum unsigned deviation δ_{max} observed for these bands. As expected, the experimental conformers observed (in red and blue) are assigned to species found to be among the most stable conformations of their backbone category (in corresponding colors). For the sake of comparison, the same data are also given for the two observed conformers of Ac-Cys-Phe-NH_2, taken from ref. 5

Ac-Cys-Phe-NH_2[a]

Exp. IR on UV band A			3413		3281		3364		3521		-
Exp. IR on UV band B			3431		3442		3389		3521		-
Conf.	ΔG (kJ.mol^{-1})	ΔH (kJ.mol^{-1})	NH_{Cys}	NH_{Phe}	NH_2 s	NH_2 a	A $\langle\delta\rangle$	δ_{max}	B $\langle\delta\rangle$	δ_{max}	H-bond pattern
a	0.0	0.0	3429	3450	3396	3529	56	115	6	8	5^{γ}-π-10
b	6.2	4.3	3410	3278	3379	3525	6	15	50	111	5^{γ}-7-7

Ac-Sec-Phe-NH_2

Exp. IR on UV band A			3401		3281		3363		3521		-
Exp. IR on UV band B			3422		3441		3390		3522		-
Conf.	ΔG (kJ.mol^{-1})	ΔH (kJ.mol^{-1})	NH_{Sec}	NH_{Phe}	NH_2 s	NH_2 a	A $\langle\delta\rangle$	δ_{max}	B $\langle\delta\rangle$	δ_{max}	H-bond pattern
a	0.0	0.0	3419	3446	3399	3530	57	118	6	9	5^{γ}-π-10
b	2.4	7.4	3404	3461	3379	3524	51	98	13	20	5^{γ}-f-7
c	2.9	12.6	3458	3443	3386	3528	62	105	12	21	5-π-9^{γ}
d	3.0	4.1	3417	3474	3402	3529	64	121	14	33	5^{γ}-f-10
e	3.5	12.4	3440	3432	3383	3527	54	102	6	10	5-6^{γ}/π-9^{γ}
f	4.0	14.0	3469	3431	3421	3541	74	140	22	31	5-π-9^{γ}
g	5.0	13.0	3433	3427	3425	3540	65	144	17	35	5^{γ}-5-π
h	6.2	14.6	3439	3426	3439	3554	73	145	22	36	5^{γ}/π-5-f
i	6.8	6.7	3400	3303	3383	3527	12	22	43	87	5^{γ}-7-7
j	6.8	9.9	3409	3433	3389	3528	47	102	7	13	5^{γ}-π-10

[a] Experimental data taken from ref. 5; theoretical results on Ac-Cys-Phe-NH_2 recalculated from structures in ref. 5 at the RI-B97-D3(BJ)-abc level, where only conformations assigned to observed conformers have been considered.

populations of the two conformers to Ac-Sec-Phe-NH_2 observed; a discussion also relevant to Ac-Cys-Phe-NH_2.

First, one observes a significant disagreement between the apparent relative populations of conformers A and B in favor of A (Fig. 2) and the high relative energy of conformation i compared to the most stable form a, at $\sim$7 kJ mol^{-1}, whatever the temperature considered Table 1. Intensities in R2PI spectra not only depend on populations but also on photoionization efficiencies, which can be sometimes limited by excited lifetime issues.[47] But no evidence for such type of limitations could be found in previous studies of Phe-containing peptides.[46,48] If one puts aside the issue of the precision of energetic calculations (which is probably of the order of 5 kJ mol^{-1}, and could account, at least partially, for the discrepancy), explanation of the apparently out-of-equilibrium populations could also be looked for by considering the large weight of the double γ-turn basin. This latter is composed of 7_L-7_L but also 7_L-7_D and 7_D-7_L conformational families at high temperature, which are expected to funnel into the most stable form of the basin, *i.e.*, conformer i, in the late expansion.

Second, since the form i, observed as A in the expansion, turns out to be the ninth most stable calculated conformer at high temperature, one can wonder why conformations more stable than i remained unobserved. One should however notice that in the late expansion, enthalpy values (ΔH) should be the most relevant energetic parameters: in this case, conformer i now becomes the third most stable, whereas conformers c, e, f, g, h and j lie at much higher energies. This suggests that the effective temperature that describes the conformational populations could be significantly lower than 300 K, which would explain why these conformations, that feature a weak H-bonding content (no strong 7 or 10 H-bonds), and are hence entropically penalized at low temperature are not observed. Alternatively, a rationalization for the experimental absence of conformers b and d might be found in the geometrical parameters of these structures displayed in Table S1 (ESI†). d and a conformers are both β-turns, presenting therefore similar backbones. The main structure difference lies in the orientation of the Phe side chain, meaning that the isomerization barrier between both structures is probably not too high, allowing conformer d to relax towards the most stable a form. The same argument holds in the case of the b conformer, which mainly differs from i by the ψ_{Sec} dihedral angle; it is thus likely that b relaxes into the i conformer, which was found to be slightly more stable at low temperature.

3.4 Spectroscopic and structural evolution along the series of side chain chalcogen heteroatoms

From a structural point of view, both types of backbone conformations present (7-7 and 10 backbones) were found to be strikingly similar along the series of chalcogen heteroatoms from O to Se, as evidenced by the IR spectroscopy synthesis of Fig. 4. IR bands of the NH_{Phe} (C7 (A) or π (B)) and NH_2 s (C7 (A) or C10 (B)) groups, are almost superimposable, demonstrating

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

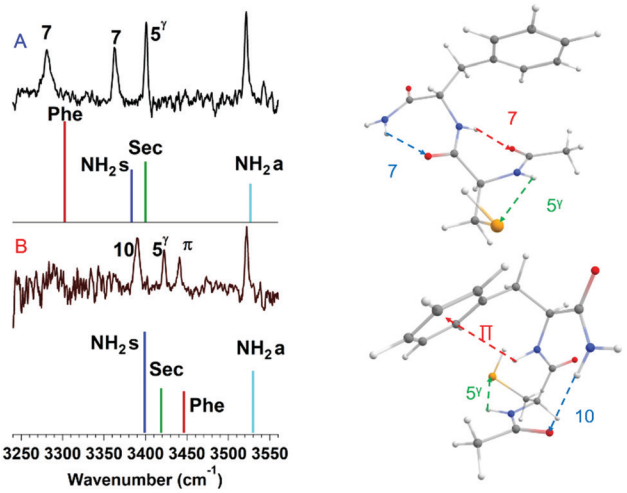

Fig. 3 IR/UV double resonance spectra of the two identified conformers in the UV spectrum of Ac-Sec-Phe-NH_2, together with simulated stick IR spectra of the conformations, which provide the best fit both in terms of frequencies (see Table 1) and intensities. The red bars are the stretching bands of the NH group localized on the Phe peptide, the blue (resp. light blue) bars are those of the symmetric (resp. antisymmetric) mode of the NH_2 group and the green bars those involving the NH group of the Sec peptide. The relevant structures (RI-B97-D3(BJ)-abc/def2-TZVPPD level of theory) along with the color-coded H-bonds present are depicted on the right panel. The present IR spectra were carried out by tuning the UV laser on an intense transition of the UV spectrum, namely at 37497 and 37581 cm^{-1} for conformer A and B respectively.

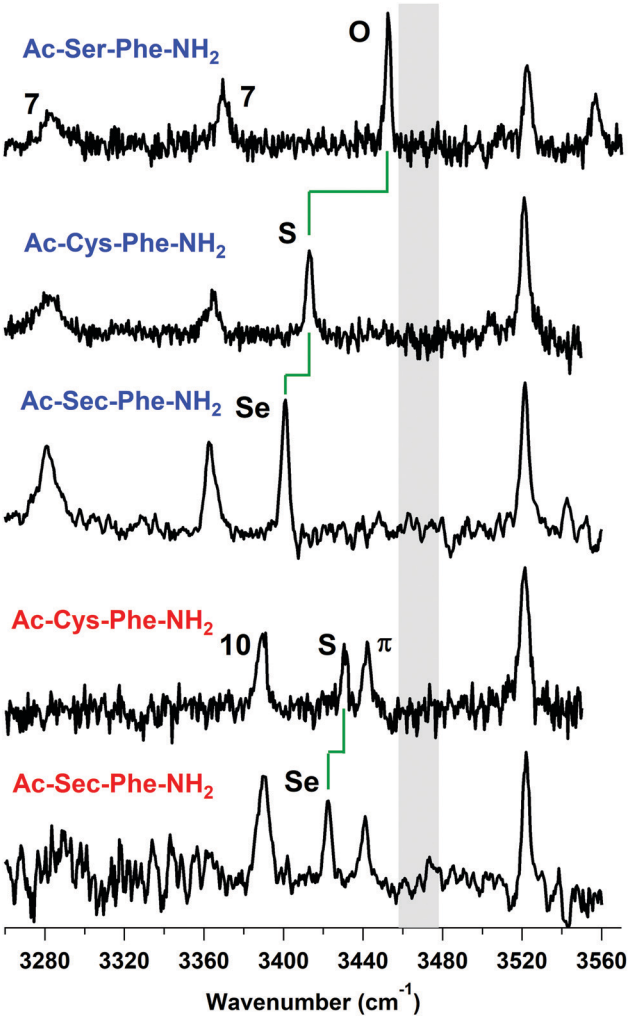

Fig. 4 Evolution of the NH stretch signatures of the 5^{γ} bands in A and B conformers, namely 7-7 (γ-turn, blue labels) and 10 (β-turn; red labels) environments respectively, along the series of side chain chalcogen heteroatoms. For reference, the free NH region is indicated by the grey bar.

that the side chain/main chain interactions do not substantially modify the backbone structure, including the main chain/main chain interactions. In contrast, the bands assigned to the NH stretches of the 5^{γ} NH···Y bonds are all red shifted relative to the free NH stretch region (grey bar in Fig. 4). In conformers A (γ-turn environment), both S and Se bands are found in the same region whereas the O band seems to be much less red shifted. In conformers B (β-turn environment), again S and Se bands exhibit similar shifts, with in both cases a slightly larger shift for selenium.

Owing to the decrease of the Y atom electroaffinity and therefore of the polar character of the YH bond along the chalcogen series, the trend of the strength of the 5^{γ} bond to increase appeared intriguing and somehow surprising. In an attempt to rationalize this evolution, a detailed theoretical analysis has been carried out at the quantum chemistry level.

4 Characterization of the local structural 5^{γ} motif of the Cys and Sec residues

4.1 Existence of an intrinsic 5^{γ} motif

Structure. The existence of an "intrinsic" 5^{γ} structural motif, formed in absence of any influence external to the residue, such as a $^{\gamma}6$ H-bond, was tested by optimizing a truncated version of the Ser, Cys and Sec residues in which the side chain can only interact with the N-terminal side amide next to it, namely the (*N*-(2-ol/thiol/selenol-1-methyl-ethyl)acetamide) model systems (see Fig. S1, ESI†). The optimized geometries, depicted in Fig. 5, show that the three side chains adopt the same type of folding motif, where the chalcogen acceptor lies close to the amide plane, with one of its lone pairs facing the amide NH bond. A close examination of the structure (Fig. 5) and of the structural parameters (Table 2) however reveals that Ser seems to behave differently. The 5^{γ} motif of Ser exhibits the most coplanar geometry, the O atom lying close to the CNCα amide plane as judged by the CNCαY dihedral close to 180°. In the very similar Cys and Sec structures, the Y atom is off the amide plane (smaller CNCαY dihedral values), but, in contrast, the Cβ-Y and amide NH covalent bonds appear to be more coplanar (smaller HN-CβY dihedral values).

NBO analysis. In order to quantify the stabilizing role of electron delocalization through hyperconjugation effects and H-bonding on the (*N*-(2-ol/thiol/selenol-1-methyl-ethyl)acetamide) model systems, a NBO analysis was carried out and showed

 This journal is © the Owner Societies 2020

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

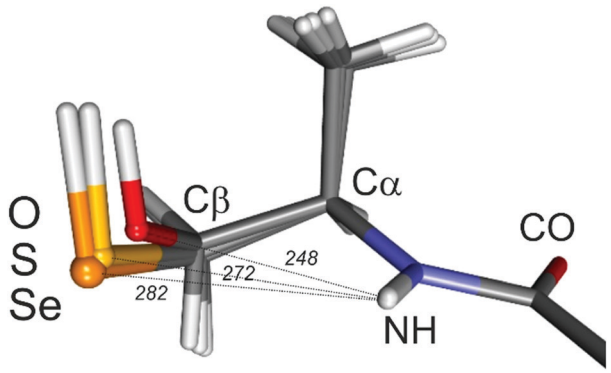

Fig. 5 Intrinsic 5^{γ} motifs from the structure optimization of the *N*-(2-ol/thiol/selenol-1-methyl-ethyl)acetamide model molecules at the RI-B97-D3(BJ)-abc/def2-TZVPPD level of theory. For the sake of comparison the amide heavy atoms N and C and the Cα atoms of the three molecules have been overlaid.

important HC effects (Table 3). The most prominent ($\sim$10 kcal mol^{-1}) is the strong $n_N \rightarrow \sigma^*_{C\alpha C_{Me}}$ HC interaction internal to the backbone, which accounts for about 50% of the total HC sum and tends to force the ϕ dihedral to adopt a 90° value, in a way similar to what was previously reported for the ϕ dihedrals of γ-turns.[40] This trend is shared by the 5^{γ} motifs all along the chalcogen series. A second HC interaction of the same type ($n_N \rightarrow \sigma^*_{C\alpha C\beta}$), but this time pertaining to the orientation of the side chain relative to the amide bond, is less intense and hence counterbalances partially the effect of the first one. Meanwhile two medium strength HC interactions pertaining to the relative orientation of the CαH and NH/CβY covalent bonds, play a major role in distinguishing the chalcogen atoms. These interactions, $\sigma_C\alpha_H \rightarrow \sigma^*_{NH}$ and $\sigma_C\alpha_H \rightarrow \sigma^*_{C\beta S/Se}$ (Fig. 6, bottom), involve the CαH bond as the donor NBO; the acceptor NBOs being the σ NBOs of the NH and CβS/Se covalent bonds. The stabilization energies arising from these interactions amount to 4.6 and 6.7 kcal mol^{-1} resp. for Cys and to 4.8 and 7.8 kcal mol^{-1} resp. for Sec (see details in Table 3). These HC interactions tend to maximize the NBO overlaps, which is achieved for NH and CβS/Se covalent bonds parallel to the CαH bond, and chiefly contributes to flatten the 5^{γ} motif (coplanar NH and CβS/Se bonds). In serine, the same interactions are present but tend to be weaker (Table 3).

Besides these main interactions, the NH$\cdots$S/Se 5^{γ} H-bonding interactions can also pretend to contribute to the motif stabilization, through $n_{S/Se} \rightarrow \sigma^*_{NH}$ and $n'_{S/Se} \rightarrow \sigma^*_{NH}$ NBO interactions (Fig. 6 top). These interactions describe the delocalization from the S/Se non-bonding NBOs, labelled n and n′, corresponding to the S/Se lone pairs, towards the anti-bonding NBO of the NH bond; the NBO oriented perpendicular to the CβS/SeH plane ($n'_{S/Se}$) being the major contributor (in red Fig. 6; also see details in Table 3). The strength of this delocalization-driven interaction, assessed from the ΣE_{HB} indicator as previously shown,[34–36,40,49] indicates a weak stabilization energy interaction, in the kcal mol^{-1} range, *i.e.*, in the same energy range than the intraresidue C5 interaction between amide groups in extended backbones. This demonstrates the ancillary nature of this interaction, which resembles a H-bond but is non linear and much weaker than regular ones (*e.g.*, 7 or 10 H-bonds,[40] rather found in the 3–10 kcal mol^{-1} range[40]). In contrast, 5^{γ} H-bonding in Ser appears nearly inexistent, with a stabilization energy of 0.2 kcal mol^{-1}, *i.e.*, weaker than in Cys and Sec by a factor of $\sim$5.

Discussion. The picture arising from this analysis features strong HC interactions, controling the orientation of the side chain relative to the backbone. The HB interaction adapts to this strong constraint and remains weak due to this frustration. The details of the constraints, however, depend upon the chalcogen atom. Strong in the case of S and Se, the HC forces the NH and CβY bonds to remain roughly parallel, leading to an off-amide plane position of the S and Se atoms, and thus to an elongated H-bond. In contrast, this does not happen in serine due to weaker HC constraints: the oxygen atom lies very close to the amide plane. The strength of the 5^{γ} H-bond allowed by

Table 2 Structural parameters relevant to the description of the intrinsic 5^{γ} motif of Ser, Cys and Sec, and their counterparts observed in the γ-turn environment for Ser and in the γ- and β-turn environments for Cys and Sec. ϕ is the C_{CO}–N–Cα–C_{Me} dihedral equivalent to the first Ramanchandran dihedral of α-peptides. The C_{amide}–N_{amide-}-Cα–Y and H_{amide}-N_{amide}-Cβ-Y dihedrals are used to assess the position of the Y atom above the amide plane and the coplanarity of the amide NH and Cβ-Y covalent bonds, respectively. H-bond distances are given in pm. For the sake of comparison, the NH$\cdots$Y distances in the intermolecular H-bonded complexes of *trans*-methylacetamide (CH_3–CONH–CH_3) with Me–YH, Y = O, S and Se are also given

	O[a]					S					Se				
Y	ϕ	CN-CαO	HN-CβO	NH$\cdots$O	OH$\cdots$OC	ϕ	CN-CαS	HN-CβS	NH$\cdots$S	SH$\cdots$OC	ϕ	CN-CαSe	HN-CβSe	NH$\cdots$Se	SeH$\cdots$OC
H-bond				5^{γ}	$^{\gamma}6$				5^{γ}	$^{\gamma}6$				5^{γ}	$^{\gamma}6$
Intrinsic 5^{γ} NH$\cdots$Y	**−84°**	**177°**	**19°**	**248**	—	**−89°**	**173°**	**10°**	**272**	—	**−92°**	**171°**	**8°**	**282**	—
In γ-turn env. (5^{γ}-7-7; A conf.)	−82°	178°	24°	246	212	−82°	180°	15°	259	239	−83°	179°	15°	266	251
In β-turn env. (5^{γ}-π-10; B conf.)	—	—	—	—	—	−72°	−168°	26°	265	333	−70°	−166°	26°	274	338
Intermolecular NH$\cdots$Y *trans*-methyl-acetamide$\cdots$$CH_3$-YH			—	204				—	254				—	266	

[a] For Ser, only conformation assigned to observed conformers of ref. 5, *i.e.*, the γ-turn has been considered and reoptimized at the RI-B97-D3(BJ)-abc/def2-TZVPPD level.

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

Table 3 Relevant second order stabilization energies $E(2)$, given in kcal mol^{-1}, obtained from the NBO analysis of the intrinsic 5^{γ} motifs and their counterparts observed in γ- and β-turn environments: for the 5^{γ} and $^{\gamma}6$ H-bonds, $E(2)$ energies arising from the interaction between several NBO donors (n_Y and n'_Y) and (n_O, n'_O and π_{CO}) and the corresponding acceptor NBOs (σ^*_{NH} and σ^*_{YH} resp.); for hyperconjugation effects, interactions between the amide nitrogen lone pair n_N or the $\sigma_{C\alpha H}$ NBO and several vicinal covalent bound σ NBOs

	Ser		Cys			Sec		
	Intrinsic	γ-turn	Intrinsic	γ-turn	β-turn	Intrinsic	γ-turn	β-turn
	5^{γ} model	5^{γ}-7-7 A	5^{γ} model	5^{γ}-7-7 A	5^{γ}-π-10 B	5^{γ} model	5^{γ}-7-7 A	5^{γ}-π-10 B
5^{γ}								
$E(2)$ $n_Y \rightarrow \sigma^*_{NH}$	<0.05	<0.05	0.1	0.3	0.2	0.1	0.3	0.2
$E(2)$ $n'_Y \rightarrow \sigma^*_{NH}$	0.2	0.3	0.9	2.1	1.2	1.0	2.6	1.5
Sum of the $E(2)$	0.2	0.3	1.0	2.4	1.4	1.1	2.9	1.7
$^{\gamma}6$								
$E(2)$ $n_O \rightarrow \sigma^*_{YH}$	—	0.3	—	0.3	—	—	0.3	—
$E(2)$ $n'_O \rightarrow \sigma^*_{YH}$	—	1.8	—	0.6	—	—	0.3	—
$E(2)$ $\pi_{CO} \rightarrow \sigma^*_{YH}$	—	0.4	—	0.4	0.1	—	0.3	0.1
ΣE_{HB}: sum of $E(2)$	—	2.5	—	1.3	0.1	—	0.9	0.1
Hyperconjugation								
$E(2)$ $n_N \rightarrow \sigma^*_{C\alpha C_{Me}}$	9.5	11.1	9.6	11.9	12.9	9.5	12.1	13.0
$E(2)$ $n_N \rightarrow \sigma^*_{C\alpha C\beta}$	2.2	1.6	3.2	1.8	0.9	**3.6**	1.9	0.8
$E(2)$ $\sigma_{C\alpha H} \rightarrow \sigma^*_{NH}$	3.9	3.3	4.6	3.6	3.3	4.8	3.7	3.4
$E(2)$ $\sigma_{C\alpha H} \rightarrow \sigma^*_{C\beta Y}$	5.3	4.3	6.7	5.6	6.0	7.8	6.6	7.0
Sum of $E(2)$	20.9	20.3	24.1	22.9	23.1	25.7	24.3	24.2

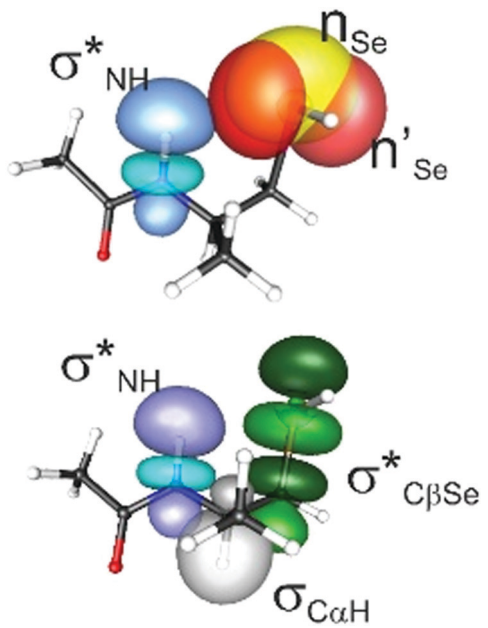

Fig. 6 Relevant Natural Bond Orbitals describing the H-bonding and hyperconjugative interactions in the *N*-(2-selenol-1-methyl-ethyl)acetamide molecule, a minimalist model for the 5^{γ} motif. Bottom: NBOs relevant to the description of the hyperconjugative interactions involving the CαH bond (donor σ NBO in grey) and its vicinal neighbors NH and CβSe (acceptor σ* NBOs in blue and dark green respectively). Top: NBO interactions relevant to H-bonding: in blue the σ^*_{NH} accepting NBO overlaps partially the donor NBOs associated to the n_{Se} and n'_{Se} lone pairs (in green and red resp.); NBOs contours are shown at a density of 0.05 u.a. See text for details.

these constrains depends upon the features of the chalcogen atom. S and Se give rise to a weak but significant HB interaction, with a slight advantage to the Se H-bond. The serine case is obviously much more hampered, especially because of the small van der Waals radius. In this respect, one should notice that the H-bond strength is multifactorial. The interaction strength is expected to differ among the acceptor O/S/Se centers, because (i) the structural preferences imposed by the HC depend upon the Y atom, (ii) the position of the chalcogen relative to the Cα atom depends on the length of the Cβ-Y covalent bond and hence upon the covalent radii of the Y atom, (iii) electronic factors vary among the chalcogens, in particular, the atomic van der Waals radii, the polar character of the YH bond, *etc.* All these factors will simultaneously affect the ability to establish a H-bond, and the HB strength observed *in fine* will result from a compromise, which might not be amenable to a simple rationalization, but is nevertheless properly modelled by quantum chemistry, as testified by the good agreement between experimental and theoretical NH stretch frequencies reported in Section 3. From this theoretical modelling, one can conclude that the heavier the chalcogen, the stronger the H-bonding in the intrinsic 5^{γ} motif.

4.2 The 5^{γ} motif in secondary structure environments

The aim of this sectioin is to examine how external constraints imposed by the rest of the molecule can influence the 5^{γ} motif. Fig. 7 shows the 5^{γ} motifs when embedded in secondary structures as in the dipeptides studied. The overall backbone secondary structure imposes its own constraints, in particular it controls the ϕ dihedral of the residue (which barely varies with Y; see Table 2), as illustrated by the similarities in the backbone 7 or 10 H-bonds of Fig. 4. In contrast with Fig. 5, this backbone environments also allow the formation of $^{\gamma}6$ H-bonds between the side chain and the next amide group of the backbone.

 This journal is © the Owner Societies 2020

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

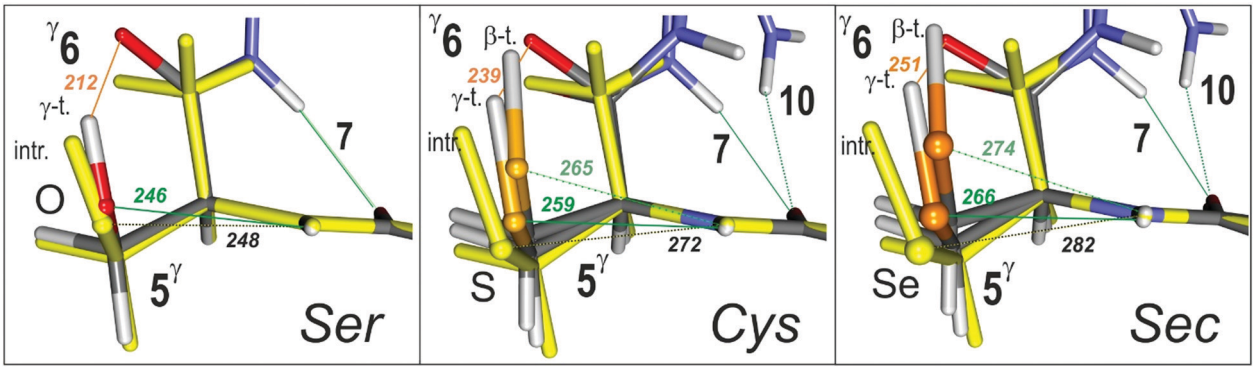

Fig. 7 Comparison of the intrinsic 5^{γ} motifs with their counterparts observed in the γ-turn environment for the Ser residue (left panel) and γ- and β-turn environments, for the Cys and Sec residues (center and right panel resp.), as obtained by quantum chemistry calculations. 5^{γ} and $^{\gamma}6$ H-bond distances are given in pm. For the sake of clarity the backbone 7 and 10 H-bonds of the secondary structures are also shown.

These constraints add themselves to the intrinsic ones of the side chain, but do not necessarily act in the same direction. This is indicated for instance by the slight decrease of the HCs internal to the 5^{γ} motif (bottom line, Table 3). Eventually, the 5^{γ} motif has to adapt itself to the resultant of these constraints. In the Cys and Sec cases, the motif experiences significant distortions (HN-CβS dihedral angles in Table 2). In the γ-turn, the Y atoms get closer to the amide plane (CNCαY dihedral angles ~180°) and the HB distances decrease by 13 and 14 pm compared to the respective intrinsic 5^{γ} motifs (Fig. 7). In the β-turn environments, the distortion increases further, the Y atom deviates again from the amide plane and concomitantly the HB distance increases again, by 6 and 8 pm respectively. No such variations are observed with Ser, because the structure imposed by the γ-turn is similar to that of the Ser intrinsic motif, also indicating a nearly inexistent H-bonding in the γ-turn (Table 3). These trends are corroborated by the NBO analysis of the 5^{γ} H-bonds of the dipeptides studied (Table 3). In fact, for Cys and Sec the γ-turns exhibit the shortest HB distances and the largest stabilization energies (in the 2–3 kcal mol^{-1} range). The β-turns show significantly weaker stabilizations, intermediate between those of the γ-turn environment and the intrinsic 5^{γ} motif. In contrast, with Ser, the γ-turns exhibit no 5^{γ} H-bonds, *i.e.* only a very weak stabilization energy (~0.3 kcal mol^{-1}) as in the intrinsic 5^{γ} motif.

Compared to the isolated 5^{γ} motif, the presence of the peptide backbone environment also allows the formation of YH···O intraresidue $^{\gamma}6$ interactions, where the side chain acts as a donor, depending upon the backbone secondary structure. The atoms involved in the H-bond being the same, the interatomic YH···O distance can be used as a direct indicator of the bond strength (Fig. 7 and Table 3). The $^{\gamma}6$ interaction is quite significant (2.5 kcal mol^{-1}) in conformer A of Ser with a short OH···O interatomic distance of 212 pm and a 5^{γ} motif similar to the intrinsic one, suggesting a mutual compatibility of the $^{\gamma}6$ H-bond and of the distorted 5^{γ} motif. The $^{\gamma}6$ interaction tends to vanish when going to Cys and Sec with strongly elongated YH···O distances (239 and 251 pm) and decreasing NBO energies (1.2 and 0.9 kcal mol^{-1} in Cys and Sec respectively). In conformers B, the NBO HB strength is found to be below the NBO interaction detection limit (0.05 kJ mol^{-1}), in line with the large S/SeH···OC distances calculated (Table 3), so that H-bonds can no longer be considered.

Spectroscopy and cooperative effects. The simultaneous presence of NH···Y 5^{γ} and YH···O $^{\gamma}6$ interactions can lead to connected H-bonds (Fig. 1 right panel), when the peptide bonds that surround the Cys/Sec residue are donor and/or acceptor, potentially yielding to cooperative effects. In order to assess these effects, the *N*-(2-thiol/selenol-1-methyl-ethyl)acetamide model molecules, already considered for the intrinsic motifs, were again considered. Starting from the geometry of the optimized dipeptides, the system was truncated at the level of the central amide carbonyl group; this latter being substituted by a methyl group, which was then optimized, the rest of the model remaining frozen at the dipeptide geometry (Fig. S1, ESI†). The results (Tables S2 and S3, ESI†) show that both the NBO stabilization energies (ΣE_{HB} indicator; Table S2, ESI†) and the NH stretch frequencies of these truncated molecules (Table S3, ESI†) are very close to their counterpart in the corresponding Ac-Cys-Phe-NH_2 minima (within less than 10% for the NBO energies and a few wavenumbers for the frequencies). This demonstrates the absence of significant cooperative electronic effect from the acceptor character of the CO site of the amide involved in the 5^{γ} bond, or the presence of a YH···OC $^{\gamma}6$ H-bond. Furthermore, this highlights that the NH stretch frequency is a relevant indicator for the H-bonding in the 5^{γ} motif.

NBO *vs.* spectroscopy. In order to further test this view, the experimental frequencies along the series of 5^{γ} H-bonds for Ser, Sec and Cys, in the β- and γ-turn environments have been correlated to the ΣE_{HB} NBO indicators (Table S3 and Fig. 8). Owing to the good agreements between calculated and experimental values, the corrected frequencies calculated for the intrinsic 5^{γ} models have also been added for the sake of completeness. The correlation obtained is well described by a linear dependence of the shift with ΣE_{HB}. The slope obtained, of −25 cm^{-1}/(kcal mol^{-1}), is in line with that found for the same correlation within a set of weak amide–amide NH···O interactions (−28 cm^{-1}/(kcal mol^{-1})),[40] having all the same

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

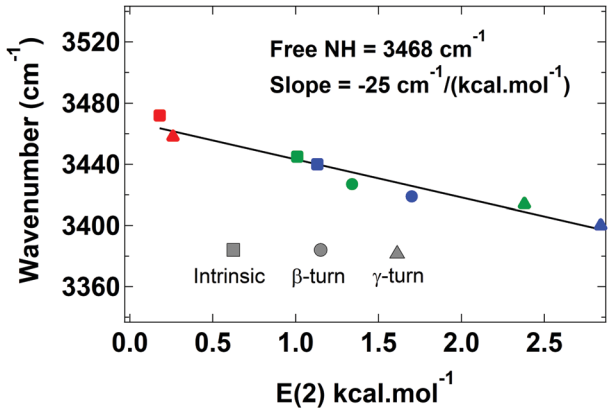

Fig. 8 Experimental stretching frequencies of 5^{γ} bonded along the Ac-Xxx-Phe-NH_2 (Xxx = Ser, Cys, Sec) series of molecules as a function of the NBO stabilization indicator (ΣE_{HB}), for several backbone environments (triangles: in γ-turn; disks in β-turns) completed by the three calculated intrinsic motifs (squares). The nature of the residue is color-coded (Ser, red; Cys, green; Sec, blue). The linear fit (solid line) gives an intercept of 3469 cm^{-1} and a slope of −25 cm^{-1}/(kcal mol^{-1}).

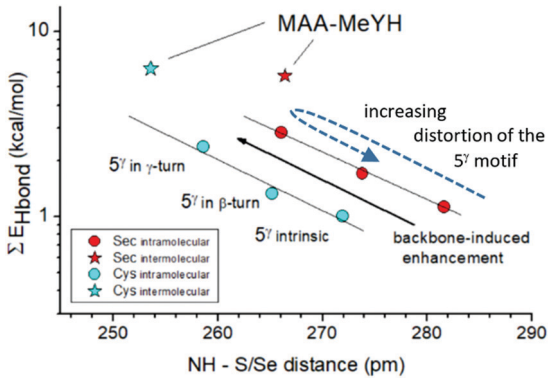

Fig. 9 Semi-log plot of the sum of the HB stabilization energies (ΣE_{HB}), as obtained from the NBO analysis of the Cys and Sec series of 5^{γ} H-bonds (in the intrinsic model, in γ- and β-turn backbone environments, disks) and of the *trans*-methylacetamide-CH_3-YH complex (Y = S and Se, stars); see details in Table S3 (ESI†). Structures sharing the same approach are expected to be aligned (see ref. 40). A backbone-induced shrinkage of the H-bond causes a bond strength increase (black arrow).

donor and acceptor but oriented according to different geometrical approaches. This is inasmuch noticeable as the nature of the proton acceptor changes along the present series, suggesting a universal validity for the ΣE_{HB} indicator. One can also notice that the intercept obtained from the present linear dependence at 3468 cm^{-1} provides a reference shift for the 5^{γ} H-bonds considered. This fixes the issue of such a shift reference, which is otherwise tricky because systems presenting a free NH stretching mode in ref. 40 have their frequency spread over 40 cm^{-1} depending upon the secondary structure in which the residue is embedded. This was interpreted in terms of subtle hyperconjugation effects between the NH bond considered and vicinal neighbor covalent bonds (see for example ref. 3 and 40).

Constrained H-bonding

The constraints applied on the 5^{γ} H-bonds in γ- and β-turns, has been analyzed following a methodology already successfully applied to constrained amide–amide interactions.[40] For this purpose, the dependence of the NBO indicator has been plotted as a function of the H-bonding distance (Fig. 9).

Let us first consider the series of Cys 5^{γ} intramolecular bonds. The three Cys 5^{γ} H-bonds (intrinsic, γ- and β-turn) follow a roughly linear trend (Fig. 9, cyan disks), reminiscent of what is observed for sets of amide–amide interactions that share a same H-bonding approach (*e.g.*, 5 amide–amide bonds in ref. 40). This fairly agrees with the similar geometrical approaches calculated for these side chain structures as illustrated in Fig. 7. The same features are found for the Sec 5^{γ} series, as shown by the red disks of Fig. 9.

The linear dependence of Fig. 9 is interpreted in terms of existence of constraints external to the residue and exerted by its environment, which can either elongate or shrink the H-bond.[40] The observed 5^{γ} bonds are stronger than the intrinsic 5^{γ} feature, with smaller HB distances. This demonstrates that the secondary structures have a compressing, and thus enhancing, effect on the 5^{γ} bond, in qualitative agreement with the weaker HC interactions (which stabilize the 5^{γ} motif) in these environments, as discussed in the previous section. More precisely whereas the external constraints and structural distortions of the 5^{γ} increase along the intrinsic, γ-turn and β-turn series (see Fig. 7 and Table 2), the HB distance first decreases and then increases again (dotted arrow in Fig. 9). With sulfur, the minimum HB distance (259 pm in the γ-turn) gets close to the unconstraint intermolecular H-bonded complex *trans*-methylacetamide···CH_3SH in the (254 pm). With selenium, the 5^{γ} distance in γ-turn (265 pm) is even shorter than the intermolecular H-bond (266 pm), although the H-bond is typically twice weaker than in the complex. Regarding now the comparison between S and Se, the difference in H-bonding distances between the two systems is illustrative. The difference in intrinsic 5^{γ} distance between Cys and Sec (10 pm) is smaller than the difference in intermolecular HB (13 pm) already suggesting a significantly stronger intrinsic 5^{γ} motif with selenium. In β- and γ-turns, this difference tends to further decrease, confirming a larger propensity of the selenium atom to establish a stronger and hence shorter HB under constraining environments.

5 Conclusions

The present gas phase analysis documents the existence of an intraresidue H-bonding motif, labelled 5^{γ}, for cysteine and selenocysteine residues, which bridges the side chain, where the S/Se atom plays an acceptor role, and the NH bond of the same residue, in a non-linear H-bonding geometry. NBO analysis on minimalist models, namely in absence of external constraint, provide evidence for the intrinsic character of such a 5^{γ} motif, which leads to a significant stabilization for Cys and Sec, in contrast to what could be expected from the evolution of the YH bond polar character along the chalcogen series. Comparison with serine also emphasizes a significant

 This journal is © the Owner Societies 2020

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

difference in nature. Whereas the Ser 5^{γ} motif is essentially stabilized by hyperconjugation effects, with a negligible H-bonding, those of Sec and, to a lesser extent, Cys are stabilized by both a specific hyperconjugation effect involving the chalcogen atom (resulting in a departure of the Y atom from the donor amide plane) and a significant NH···Y hydrogen bonding. The strength of these interactions is controlled by various atomic features of the chalcogen atom, among them its van der Waals and covalent radii. In such a constrained geometry context the O atom, in contrast to Se and to a lesser extent S, is too small to develop significant 5^{γ} H-bonding, whereas the donor property of the hydroxy group favours the formation of a $^{\gamma}6$ OH···O bond instead.

When embedded in secondary structures like in a γ-turn or in the first residue of a β-turn, the external constraints imposed by the backbone can compete with the HC effects and the NH···Y H-bond adapts itself to this new environment, giving rise to stronger H-bonding interactions, depending upon the type of secondary structure. In the present case, the γ-turn provides more favorable geometrical conditions for the H-bonding than a β-turn, in particular shorter 5^{γ} H-bond distances. The evolution of the spectroscopic IR signatures in the NH stretch region of the weak 5^{γ} H-bonds along the Ser, Cys and Sec series, presently reported, reflects a compromise between the several factors mentioned above and illustrates the H-bond enhancements due to the strong constraints imposed by the secondary structure.

In contrast to what was observed for asparagine,[7] the present local NH···S/Se H-bonds in these motifs are too weak to significantly influence by themselves the α-peptide backbone structure. Their properties, however, strongly depend upon the side chain structure together with the type of peptide considered (α-, β-, *etc.*): they are currently exploited in the field of bioinspired foldamers to stabilize unusual secondary structures of synthetic peptides.[50]

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

Support from the French National Research Agency (ANR) is acknowledged (Grants ANR-17-CE29-0008 ''TUNIFOLD-S'' and ANR-16-CE29-0017 ''IONPAIRS''). 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. We also acknowledge the use of the computing facility cluster MésoLUM of the LUMAT federation (FR LUMAT 2764).

References

1 M. S. de Vries and P. Hobza, *Annu. Rev. Phys. Chem.*, 2007, **58**, 585–612.
2 M. Gerhards, in *Spectroscopy of Neutral Peptides in the Gas Phase: Structure, Reactivity, Microsolvation, Molecular Recognition*, John Wiley & Sons, Ltd, 2006, ch. 1, pp. 1–61.
3 E. Gloaguen and M. Mons, *Top. Curr. Chem.*, 2015, **364**, 225–270.
4 B. Yan, S. Jaeqx, W. J. van der Zande and A. M. Rijs, *Phys. Chem. Chem. Phys.*, 2014, **16**, 10770–10778.
5 M. Alauddin, H. S. Biswal, E. Gloaguen and M. Mons, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2169–2178.
6 W. Y. Sohn, S. Habka, E. Gloaguen and M. Mons, *Phys. Chem. Chem. Phys.*, 2017, **19**, 17128–17142.
7 S. Habka, W. Y. Sohn, V. Vaquero-Vara, M. Géléoc, B. Tardivel, V. Brenner, E. Gloaguen and M. Mons, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3411–3423.
8 K. N. Blodgett, J. L. Fischer, J. Lee, S. H. Choi and T. S. Zwier, *J. Phys. Chem. A*, 2018, **122**, 8762–8775.
9 P. S. Walsh, J. C. Dean, C. McBurney, H. Kang, S. H. Gellman and T. S. Zwier, *Phys. Chem. Chem. Phys.*, 2016, **18**, 11306–11322.
10 P. S. Walsh, K. N. Blodgett, C. McBurney, S. H. Gellman and T. S. Zwier, *Angew. Chem., Int. Ed.*, 2016, **55**, 14618–14622.
11 A. Bhattacherjee and S. Wategaonkar, *J. Phys. Chem. A*, 2017, **121**, 8815–8824.
12 H. S. Biswal and S. Wategaonkar, *J. Phys. Chem. A*, 2009, **113**, 12763–12773.
13 I. Chambers, J. Frampton, P. Goldfarb, N. Affara, W. McBain and P. R. Harrison, *EMBO J.*, 1986, **5**, 1221–1227.
14 F. Zinoni, A. Birkmann, T. C. Stadtman and A. Böck, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 4650–4654.
15 I. Chambers and P. R. Harrison, *Trends Biochem. Sci.*, 1987, **12**, 255–256.
16 A. Böck, K. Forchhammer, J. Heider, W. Leinfelder, G. Sawers, B. Veprek and F. Zinoni, *Mol. Microbiol.*, 1991, **5**, 515–520.
17 T. C. Stadtman, *Annu. Rev. Biochem.*, 1996, **65**, 83–100.
18 D. L. Hatfield and V. N. Gladyshev, *Mol. Cell. Biol.*, 2002, **22**, 3565–3576.
19 L. Johansson, G. Gafvelin and E. S. J. Arnér, *Biochim. Biophys. Acta*, 2005, **1726**, 1–13.
20 V. R. Mundlapati, S. Ghosh, A. Bhattacherjee, P. Tiwari and H. S. Biswal, *J. Phys. Chem. Lett.*, 2015, **6**, 1385–1389.
21 K. K. Mishra, S. K. Singh, P. Ghosh, D. Ghosh and A. Das, *Phys. Chem. Chem. Phys.*, 2017, **19**, 24179–24187.
22 A. Chand and H. S. Biswal, *J. Indian Inst. Sci.*, 2020, **100**, 77–100.
23 E. Gloaguen, H. Valdes, F. Pagliarulo, R. Pollet, B. Tardivel, P. Hobza, F. Piuzzi and M. Mons, *J. Phys. Chem. A*, 2010, **114**, 2973–2982.
24 A. M. Rijs and J. Oomens, *Top. Curr. Chem.*, 2015, **364**, 1–42.
25 *Schrödinger Release 2019-3: MacroModel*, Schrödinger, LLC, New York, NY, 2019.
26 J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J. Comput. Chem.*, 2005, **26**, 1752–2982.

Published on 01 September 2020. Downloaded by CEA Saclay on 2/15/2024 9:42:50 AM.

27 TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007, http://www.turbomole.com.
28 A. Schafer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829–5835.
29 D. Rappoport and F. Furche, *J. Chem. Phys.*, 2010, **133**, 11.
30 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
31 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**, 11860–11863.
32 M. Sierka, A. Hogekamp and R. Ahlrichs, *J. Chem. Phys.*, 2003, **118**, 9136.
33 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **242**, 652.
34 A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899–926.
35 E. D. Glendening, C. R. Landis and F. Weinhold, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 1–42.
36 I. V. Alabugin, G. dos Pasos Gomes and M. A. Abdo, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2019, **9**, e1389.
37 F. Weinhold, *J. Comput. Chem.*, 2012, **33**, 2363–2379.
38 F. Weinhold, A. E. Reed, J. E. Carpenter and E. D. Glendening, *NBO Version 3.1*.
39 M. J. Frisch, G. W. Truks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision B.01.*, 2010.
40 V. Brenner, E. Gloaguen and M. Mons, *Phys. Chem. Chem. Phys.*, 2019, **21**, 24601–24619.
41 E. Gloaguen, V. Brenner, M. Alauddin, B. Tardivel, M. Mons, A. Zehnacker-Rentien, V. Declerck and D. J. Aitken, *Angew. Chem., Int. Ed.*, 2014, **53**, 13756–13759.
42 W. Chin, F. Piuzzi, I. Dimicoli and M. Mons, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1033–1048.
43 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207–2216.
44 K. Guruprasad and S. Rajkuma, *J. Biosci.*, 2000, **25**, 143–156.
45 W. Y. Sohn, V. Brenner, E. Gloaguen and M. Mons, *Phys. Chem. Chem. Phys.*, 2016, **18**, 29969–29978.
46 Y. Loquais, E. Gloaguen, S. Habka, V. Vaquero-Vara, V. Brenner, B. Tardivel and M. Mons, *J. Phys. Chem. A*, 2015, **119**, 5932–5941.
47 M. Mons, F. Piuzzi, I. Dimicoli, L. Gorb and J. Leszczynski, *J. Phys. Chem. A*, 2006, **110**, 10921–10924.
48 Y. Loquais, E. Gloaguen, M. Alauddin, V. Brenner, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2014, **16**, 22192–22200.
49 S. Scheiner, *ChemPhysChem*, 2016, **17**, 2263–2271.
50 Z. Imani, V. R. Mundlapati, G. Goldsztejn, V. Brenner, E. Gloaguen, R. Guillot, J.-P. Baltaze, K. L. Barbu-Debus, S. Robin, A. Zehnacker, M. Mons and D. J. Aitken, *Chem. Sci.*, 2020, **11**, 9191.

 This journal is © the Owner Societies 2020