Intra-residue interactions in proteins: interplay between Serine or Cysteine side chains and backbone conformations, revealed by laser spectroscopy of isolated model peptides

ARTICLE *in* PHYSICAL CHEMISTRY CHEMICAL PHYSICS · NOVEMBER 2014

Impact Factor: 4.49 · DOI: 10.1039/C4CP04449E

4 AUTHORS:

Md Alauddin
University of Saskatchewan
21 PUBLICATIONS **27** CITATIONS
SEE PROFILE

Himansu Sekhar Biswal
The National Institute of Science Education…
32 PUBLICATIONS **397** CITATIONS
SEE PROFILE

Eric Gloaguen
French National Centre for Scientific Resea…
64 PUBLICATIONS **408** CITATIONS
SEE PROFILE

Michel Mons
Atomic Energy and Alternative Energies Co…
130 PUBLICATIONS **3,154** CITATIONS
SEE PROFILE

PCCP

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c4cp04449e

Intra-residue interactions in proteins: interplay between serine or cysteine side chains and backbone conformations, revealed by laser spectroscopy of isolated model peptides

Mohammad Alauddin,†[ab] Himansu S. Biswal,[c] Eric Gloaguen[ba] and Michel Mons*[ab]

Received 1st October 2014,
Accepted 24th November 2014

DOI: 10.1039/c4cp04449e

www.rsc.org/pccp

Intra-residue interactions play an important role in proteins by influencing local folding of the backbone. Taking advantage of the capability of gas phase experiments to provide relevant information on the intrinsic H-bonding pattern of isolated peptide chains, the intra-residue interactions of serine and cysteine residues, *i.e.*, OH/SH···OC(i) C6 and NH(i)···O/S C5 interactions in Ser/Cys residues, are probed by laser spectroscopy of isolated peptides. The strength of these local side chain–main chain interactions, elegantly documented from their IR spectral features for well-defined conformations of the main chain, demonstrates that a subtle competition exists between the two types of intra-residue bond: the C6 H-bond is the major interaction with Ser, in contrast to Cys where C5 interaction takes over. The restricted number of conformers observed in the gas phase experiment with Ser compared to Cys (where both extended and folded forms are observed) also suggests a significant mediation role of these intra-residue interactions on the competition between the several main chain folding patterns.

Introduction

The main chains of the proteins of the living world are essentially built-up from the condensation of a series of α-aminoacids in ribosomes, chosen among the famous 20 building blocks coded by DNA. This diversity among these building blocks, which differ by their side chains, ensures a wealth of possible interactions with their environment, either with the main chain of the protein or with other side chains, *in fine* allowing the protein to fold itself in a unique and specific way. Of a special interest are those aminoacids that possess (under physiological conditions) a neutral side chain with both acceptor and donor H-bonding properties, like those featuring hydroxyl/sulfhydryl groups, such as serine (Ser) or cysteine (Cys).[1–5] The side chain of these aminoacids can therefore link itself to the CO acceptor and NH donor sites of the main chain, in particular to those very close. With the short side chain of serine and cysteine, this potentially gives rise (providing a favorable main chain folding) to a local main-chain–side chain H-bonding, the side chain being donor to the carbonyl group or acceptor to the NH *of the same residue* (Fig. 1). Such an intra-residue H-bonding is especially interesting in the prospective of folding, since they can favor specific local main chain conformation, in particular by restricting the set of possible φ and ψ dihedral angles in the Ramachandran map of the residue. As a matter of fact, data mining carried out by biochemists on X-ray diffraction data strongly supports this view: for Ser, the presence of local C6 H-bonds, with the OH group engaged in a H-bond forming a 6-membered ring with the nearby CO group, is reported for a series of proteins,[4] with a relatively high prevalence. For free Cys (not bound to a metal ion), systematic investigations of the interactions have been carried

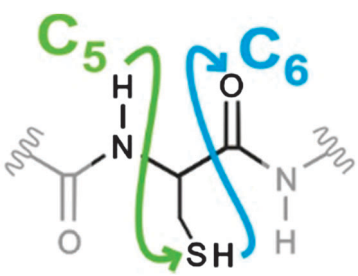

Fig. 1 Intraresidue C5 and C6 interactions shown on a cysteine residue.

[a] CEA, IRAMIS, Laboratoire Interactions, Dynamique et Lasers, CEA Saclay, Bât 522, 91191 Gif-sur-Yvette, France. E-mail: michel.mons@cea.fr
[b] CNRS, INP, Laboratoire Francis Perrin, URA 2453, CEA Saclay, Bât 522, 91191 Gif-sur-Yvette, France
[c] School of Chemical Sciences, National Institute of Science Education and Research (NISER), Bhubaneswar - 751 005, Orissa, India
† Present address: Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, SK S7N 5C9, Canada.

out, focusing on specific ones depending on the authors' interest: Chakrabarti and Pal[5] found that intra-residues interactions are dominated by S-carbonyl interactions, with the S atom sitting on top of an amide C atom. They also notice the significant prevalence ($\sim$20%) of S–π interactions when Cys possesses an aromatic environment. More recently Zhou *et al.*[3] also found S–carbonyl intramolecular interactions but noticed that they are associated with S–N intraresidue interactions, characterized by ''small NH–S angles'' ($<140^\circ$) and short H–S distances ($d < 300$ pm) together with a specific side chain orientation (with a N*i*–Cα–Cβ–S dihedral angle χ_1 close to 60°). Unfortunately the authors decided not to include them in their quantitative analysis because these interactions do not obey the classical H-bond typology. As a matter of fact, the NH–S C5 interactions, which stabilize these structures and explain these close contacts, are often evoked but not documented as due to the lack of sensitivity of X-rays diffraction data to H atoms.

In this context, beyond the first approaches led on model systems containing OH, SH and/or NH donors as well as O, S acceptors,[6–25] an in-depth understanding of the local structuring properties of a main chain induced by a Ser or Cys residue can be elegantly documented at the atomic scale by gas phase spectroscopy of isolated model peptide chains.[26–29] Such an approach, consisting in isolating and cooling down neutral species in a supersonic expansion and using IR and UV laser spectroscopy to characterize their conformations, has become popular in the past decade on various model biomolecular systems[17–25] as well as peptide chains.[26–30] These IR experiments coupled to ground state high level quantum chemistry calculations enable spectroscopists to characterize the H-bonding network of these species and therefore of their folding pattern. Given these properties, isolated systems seem to be an ideal laboratory to probe the formation and the strength of local intra-residue C5 and C6 H-bonds, like those of Ser or Cys. A first spectroscopic study of two short peptide chains containing these residues has been recently published.[30] It focused essentially onto the existence of OH/SH–OC bonds, with either spectroscopic (OH) or theoretical indirect (SH) evidence. The role of the C5 NH–X interactions was not discussed in this paper[30] and remained so far not documented in the gas phase, despite the exquisite sensitivity of the vibrational NH stretch probe to their presence.

The aim of the present work is therefore to provide extensive information on the intraresidue interactions in gas phase Ser and Cys systems, in various main chain environments. For this purpose the Cys and Ser residues have been introduced in capped two-residue sequences containing the aromatic residue Phe, required for the conformer selective analysis by the IR/UV double resonance spectroscopy: FSa, FCa, SFa and CFa sequences have been considered; XYa standing for the Ac-Xxx-Yyy-NH_2 capped peptide respectively; X and Y being the standard single-letter residue notation. The assignment have been carried out by comparison with relevant reference conformations of various other sequences previously studied in our group, GFa, FAa and FFa,[26,31–34] together with quantum chemistry calculations.

Methods

Gas phase experiments have been performed using a molecular beam set-up described in detail elsewhere.[34] It features a laser-desorption set-up coupled to a pulsed valve allowing gently vaporizing and cooling down peptide molecules in the gas phase. UV and conformation-specific IR spectroscopy is carried out using state-of-the-art mass-selective resonant two-photon ionization (R2PI) spectroscopy and IR/UV double resonance experiments respectively. Spectroscopy lasers – frequency-doubled YAG (Continuum)-pumped dye laser (Lambda-Physik FL 3002) and YAG (Quantel)-pumped single stage OPO (Laserspec) – are mildly focused simultaneously on the skimmed molecular jet, in the interaction chamber of a time-of-flight mass spectrometer. A reference signal for the double resonance experiment (''IR off'' signal) is provided by a photoion bunch formed by the UV beam sent back to the molecular beam downstream to the position where both IR and UV lasers overlap.[34] As a result, IR spectra are displayed as relative depletion signals as a function of the IR photon energy in the 3 μm range, *i.e.*, in the NH or OH stretch regions.

Quantum chemistry calculations of optimized ground state structures have been carried out at the DFT-D level (RI-B97-D/TZVPP), in order to take into account dispersive interactions, ubiquitous in these systems. Theoretical IR spectra are obtained from harmonic vibrational frequencies, scaled through mode-dependent linear scaling factors.[29,35] The set of peptide conformations studied in the group so far has allowed us to anticipate a satisfactory agreement with experiment.[29,35] In particular, for well-characterized systems, observation is accounted for by the most stable forms calculated, within a Gibbs energy range of less than a few kJ mol^{-1}. Concerning the IR spectra, calculations provide a good fit to experiment, with discrepancies usually below than 20 cm^{-1}, within the series of peptides studied.[29,35] Our confidence is such as an automated procedure has been developed, based on the following quality criteria: the energetics, the average discrepancy on all the IR bands as well as the largest discrepancy within these bands. An assignment can then efficiently be proposed, providing an extensive conformational exploration has been carried out using a standard force field.[36,37]

In the present cases, explorations have been carried out on the FCa and CFa molecules, using the AMBER force field[38] with the HyperChem package.[39] Among the first 5000 structures obtained for each molecule, 89 and 101 out of them respectively, chosen for their ability to account for the number of H-bonds experimentally observed, have been optimized at the RI-B97-D/TZVPP level of theory and their harmonic frequencies obtained (TurboMole package[40]). Harmonic frequency were scaled separately for the NH, NH_2 antisymmetric and symmetric, and OH stretching modes according to a linear scaling procedure previously introduced,[29,35] of the form $\mathrm{freq}^{\mathrm{scaled}} = a \times \mathrm{freq}^{\mathrm{harm}} + b$, with improved (*a*;*b*) values of (0.92135;188.2), (0.60872;1324.1), (0.63115;1209.8) and (0.85355;433.9) respectively, obtained from consistent gas phase series of peptides and alcohols. Gibbs energy were calculated from the harmonic frequencies using the *freeh* module of the TurboMole package.[40] Satisfactorily, small

Table 1 Assignment table for Ac-Phe-Cys-NH_2 and Ac-Phe-Ser-NH_2 species, from comparison between experimental (italics) and predicted NH stretch frequencies (cm^{-1}) for the set of most stable conformations (within 8 kJ mol^{-1} at 300 K). Conformations are sorted and labelled according to their H-bonding network (5-5, 7-7, 5-7 and 10 for β-strand, double γ-turns, mixed β-strand/γ-turn and β-turn, respectively) and to their relative Gibbs energy at 300 K, expressed in kJ mol^{-1}. The detailed structure is indicated in the first column: the first three symbols show the H-bonding status of each main chain NH (5, 7, 10, π with Phe ring or S/O for donor to the side chain heteroatom); the two next ones indicate the side chain orientation (χ_1 dihedral angle) and the parenthesis stands for the accepting moiety to which the side chain OH/SH group is donating. Electronic energy (ΔE) and 0 K stability (ΔH) are also given in kJ mol^{-1}. The agreement between the three experimental set of frequencies of a conformation and the predicted values is quantified from two quality indicators: the unsigned error averaged over the several bands observed $\langle\delta\rangle$ and the largest unsigned error among the observed bands δ^{MAX}. The best agreements are indicated by bold values

Ac-Phe-Cys-NH_2

Experimental	A	*3351*	*3355.5*	*3450.5*	*3515.5*
	B	*3388.5*	*3436*	*3450.5*	*3523.5*
	C	*3387*	*3430.5*	*3434.5*	*3520.5*

		ΔE	ΔH (0 K)	ΔG (300 K)	NH_{Phe}	NH_{Cys}	NH_2 sym.	NH_2 anti.	A $\langle\delta\rangle$	δ_{max}	B $\langle\delta\rangle$	δ_{max}	C $\langle\delta\rangle$	δ_{max}
π-S-10 g+g+ (O)	**10_1**	**0.0**	**0.0**	**0.0**	**3444**	**3438**	**3395**	**3527**	36	82	**5**	**6**	**8**	**10**
π-S-10 g−g+ (O)	**10_2**	**6.9**	**5.9**	**1.6**	**3437**	**3429**	**3379**	**3523**	30	73	**8**	**13**	**4**	**8**
π-S-10 g+g+ (π)	**10_3**	**3.3**	**3.3**	**5.6**	**3447**	**3441**	**3398**	**3528**	37	86	**6**	**10**	**11**	**13**
5-π/S-7_L ag+ (O)	**5-7_1**	**1.7**	**1.1**	**2.6**	**3451**	**3373**	**3363**	**3519**	**8**	**18**	23	63	25	57
5-π/S-7_L g+g+ (O)	5-7_2	6.9	6.6	6.9	3447	3391	3365	3509	15	35	22	45	21	39
5-5/π-S aa (-)	5-5_1	12.9	10.9	6.5	3446	3399	3391	3520	22	43	11	37	13	32

Ac-Phe-Ser-NH_2

Experimental	A	*3350*	*3397*	*3448.5*	*3512*	*3549.5*					

		ΔE	ΔH (0 K)	ΔG (300 K)	NH_{Phe}	NH_{Cys}	NH_2 sym.	NH_2 anti.	OH_{Ser}	A $\langle\delta\rangle$	δ_{max}
π-O-10 g+g+ (O)	10_1	9.3	8.9	6.5	3465	3443	3399	3528	3609	37	59
π-O-10 g−g+ (O)	10_2	14.6	14.2	11.1	3455	3434	3390	3524	3610	31	60
π-O-10 g+g+ (π)	10_3	16.4	15.7	16.4	3465	3451	3400	3527	3603	37	54
5-π/O-7_L ag+ (O)	**5-7_1**	**0.00**	**0.00**	**0.00**	**3450**	**3411**	**3365**	**3516**	**3565**	**10**	**16**

average discrepancies (less than 10 cm^{-1} for individual bands) were obtained for the best fit structures, which were also all found among the most stable structures (see Tables 1 and 2). For the Ser molecules, a simplified procedure was used. Since experiment suggests that the structures are similar to those of the most abundant ones with Cys models (A conformers), only the lowest structures of each type found with Cys have been considered by substituting Ser to Cys, and optimizing them at the DFT-D level. For comparison, they have been listed in the same tables as the Cys-based molecules (Tables 1 and 2).

Experimental results and assignments

The near UV spectroscopy of the four molecules has been obtained in the origin region of the lowest ππ* transition of the phenyl ring (Fig. 2). As expected for jet-cooled molecules the spectra are structured, providing information on both the conformational content and the Franck–Condon (FC) activity of these species. The conformational assignment of each band in the UV spectrum is obtained from observation of IR/UV double resonance depletions when probing on this band.

As a matter of fact, the UV spectra are dominated by a major feature labelled A: when Phe is in first position, it shares much similarity, in terms of spectral position and FC pattern, with the A conformer of the Ac-Phe-Ala-NH_2 (FAa) molecule, which was assigned to a C7 folding on the Ala residue combined with an extended form on Phe (pattern noted 5-7).[26,32,33,41] When Phe is in second position, the similarities of the three AFa, SFa and CFa spectra are still more striking, suggesting an assignment to 2_7 ribbon, *i.e.*, two successive C7 H-bonds along the main chain (also noted 7-7).[26,31,42] Besides these major features, IR/UV double resonance depletions reveal that minor band systems are also present with Cys (FCa and CFa), in the 37 500 and 37 600 cm^{-1} respectively, *i.e.* in the spectral regions specific to β-turns (type I) with Phe in first and second position respectively. For FCa and FSa, the present spectra confirm and extend the results by Yan *et al.*,[30] with the observation of a third minor conformer (C) with Cys.

To strengthen this tentative assignment solely based on UV spectroscopy, IR/UV double resonance spectra have been recorded (Fig. 3), enabling to access the NH/OH stretch IR spectrum of each form observed in the UV spectrum. As expected, 5 and 4 bands are generally observed in this region for molecules with Ser and Cys respectively (The FCa A form with its 3 bands is an exception, discussed in the next Section). The assignment of the A forms is straightforward, by comparison with FAa and AFa spectra:[26,31,42] FSa and FCa exhibit the same band pattern, with a C5 band on Phe (at 3445 cm^{-1}) as well as H-bonded symmetric and anti-symmetric components of the NH_2 doublet below 3400 cm^{-1}

Table 2 Assignment table for Ac-Cys-Phe-NH_2 and Ac-Ser-Phe-NH_2 species, from comparison between experimental (italics) and predicted NH stretch frequencies (cm^{-1}) for the set of conformations obtained from the Cys-containing species and by substituting an O atom to S. Conformations are sorted and labelled according to their H-bonding network (5-5, 7-7, 5-7 and 10 for β-strand, double γ-turns, mixed β-strand/γ-turn and β-turn, respectively). Relative energetic data are all given in kJ mol^{-1}. Otherwise, same legend as Table 1. A2 between brackets indicates OH/SH groups donors to the second amide group of the main chain

Ac-Ser-Phe-NH_2

Experimental	A	*3282.5*	*3369.5*	*3452.5*	*3522.5*	*3557*					
		ΔE	ΔH (0 K)	ΔG (300 K)	NH_{Ser}	NH_{Phe}	NH_2 sym.	NH_2 anti.	OH_{Ser}	A $\langle\delta\rangle$	δ_{max}
O-π-10 g+g+ (A2)	10_1	6.6	5.3	5.7	3448[a]	3445[a]	3382	3522	3627	54	100
O-f-10 g+g− (π)	10_2	10.4	7.5	4.6	3451	3465	3401	3528	3581	48	119
O-π-10 g−g+ (f)	10_3	13.3	11.2	10.9	3448	3433	3373	3525	3659	47	91
O-π-10 g+g+ (f)	10_4	12.0	10.2	9.5	3459	3450	3389	3520	3633	60	107
O-7_L-7_L g+g− (O)	**7-7_1**	**0.00**	**0.00**	**0.00**	**3451**	**3292**	**3378**	**3521**	**3561**	**5**	**10**

(a) Coupled vibrations

Ac-Cys-Phe-NH_2

Experimental	A B	*3284* *3389*	*3365.5* *3431*	*3415* *3442*	*3521* *3521*						
		ΔE	ΔH (0 K)	ΔG (300 K)	NH_{Cys}	NH_{Phe}	NH_2 sym.	NH_2 anti.	A $\langle\delta\rangle$	δ_{max}	B $\langle\delta\rangle$	δ_{max}
S-π-10 g+g+ (A2)	**10_1**	**0.0**	**0.0**	**0.0**	**3422**	**3440**	**3386**	**3523**	46	102	**4**	**9**
S-f-10 g+g− (π)	10_2	7.1	5.9	2.8	3425	3468	3397	3525	58	113	11	26
S-π-10 g−g+ (f)	10_3	7.1	6.9	6.6	3427	3431	3374	3520	42	90	7	15
S-π-10 g+g+ (f)	10_4	7.9	7.9	7.2	3427	3450	3394	3526	53	110	5	8
S-7_L-7_L g+g− (O)	**7-7_1**	**2.1**	**2.9**	**2.4**	**3415**	**3284**	**3374**	**3520**	**3**	**9**	47	104
5-5/S-π aa (A2)	5-5_1	18.0	15.9	7.9	3440	3399	3413	3533	49	113	10	18

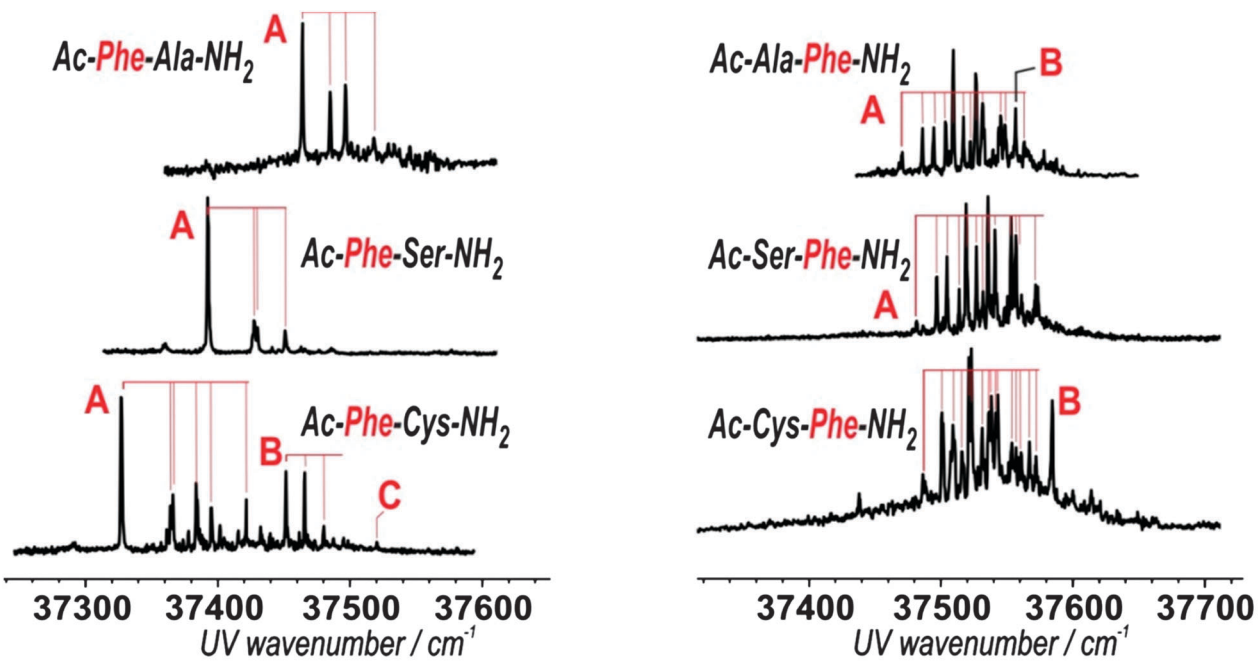

Fig. 2 Mass-selected resonant two-photon ionization spectra of the four model molecules studied in the spectral region of the first ππ* transition of the phenyl group. For comparison, the spectra of similar molecules with an alanine residue are shown. The band systems labelled with letters are identified by IR/UV double resonance (see text).

and at ~3520 cm^{-1} respectively. The only differences between both species corresponds to (i) a $NH_{Ser/Cys}$ experiencing a different environment and (ii) the H-bonded OH band of Ser at 3557 cm^{-1}. This view is confirmed by the comparison of the IR spectra

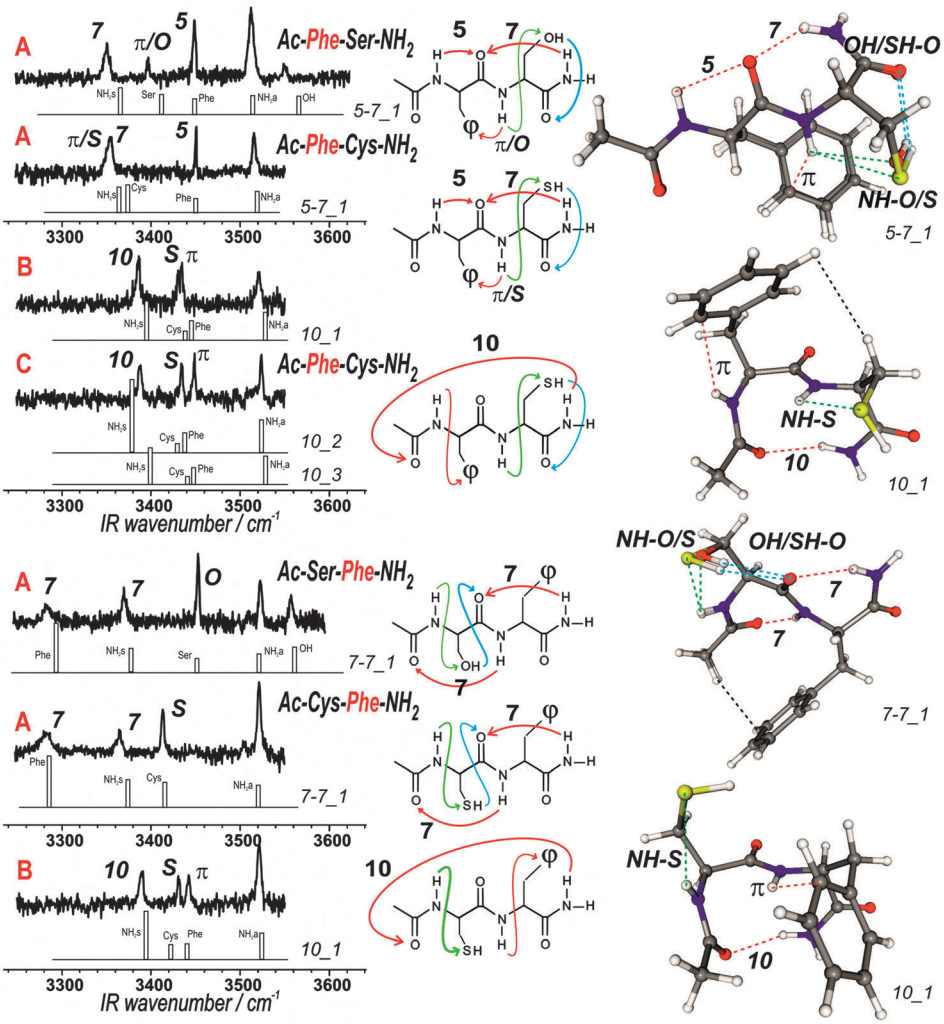

Fig. 3 IR/UV double resonance spectra (left panel) of the 7 conformers identified in the UV spectra of Fig. 2, as obtained by probing depletions on the most intense bands. Best fit calculated stick spectra (scaled RI-B97-D/TZVPP level of theory; see text) are given for each conformer. The corresponding H-bonding schemes and structures (labelled accordingly to Tables 1 and 2) are given in the central and right panels respectively. For the A conformers, the structures of Ser- and Cys-containing molecules have been displayed in such a way that backbones overlap, in order to emphasize the structural similarities.

with the calculated ones (Fig. 3) for the most stable forms (5-7_1 in Table 1), in which the 5-7 main chain H-bonding pattern is stabilized by an interaction of the $NH_{Ser/Cys}$ group with both the Phe π ring and a lone pair of O/S. In particular Fig. 3 shows that the NH_{Cys} band, apparently unobserved, probably coincides with the red-most band (symm. NH_2). The same conclusion was drawn in the previous investigation of ref. 30. When Phe comes second along the sequence, the two red-shifted bands, typical of a 2_7 ribbon, are observed in the IR spectrum, together with the H-bonded antisymmetric NH_2 band around 3520 cm^{-1}.[31,42] Again comparison with theory clearly shows that a very good agreement is matched with the most stable 2_7 ribbon form (7-7_1 in Table 2), stabilized by the OH → O H-bond on the serine side chain, as well as by the $NH_{Ser/Cys}$ interaction with the O/S lone pairs. In addition, as already noted for AFa, the large FC activity is found to be due to a large dispersive contact between the phenyl ring and the acetyl moiety, whose strength is expected to increase through ππ* excitation.

Interestingly, secondary UV bands are also observed in both species involving Cys: two in FCa and one in CFa. Their spectral location in the vicinity of that of β-turn structures, when Phe comes in first (37 500 cm^{-1}) or second position (37 560 cm^{-1}), suggests the main chains adopt β-turn structures (pattern noted 10). This is confirmed by the IR spectra, dominated in the region of H-bonded amides by isolated bands in the 3390 cm^{-1} range, characteristic of a C10 H-bond. These spectra are well reproduced by those of the most stable turns calculated: for FCa, the B and C forms can fit either of the three most stable forms (10_1 to 10_3 in Table 1). One important point to notice, however, is the significant FC activity of the B form (37 460 cm^{-1} region), which should be assigned to a close contact involving the phenyl chromophore. A basic structural examination shows that, among the three 10 structures considered, only the most stable exhibits such an interaction, between the Phe ring and the methylene group of the Cys side chain, leading to the following assignment: B to the form 10_1

and C to either of the 10_2 or 10_3 forms. It should be noted that, in contrast to conformer A, the assignments of B (and of C, not detected in ref. 30) differ from those proposed by Yan *et al.*:[30] the B form was assigned to the present 10-3 form. Our more extensive exploration of the surface, together with the use of diversified assignment arguments, leads to an alternative, overall more consistent, assignment.

For CFa B, the most stable β-turn form (10-1) provides a very good fit to the IR spectrum (Fig. 3). Owing to the substantially lesser stability and lesser frequency agreement of the next conformers (10-2 to 10-4, in Table 2), we assign the B form to this 10-1 species.

General discussion

Intraresidue interactions and energetic competition

Beyond a simple structural assignment of the conformers observed, the complete set of experimental and theoretical data on the four sequences studied provides interesting insights on the intraresidue interactions that stabilise the side chain conformation as well as the backbone conformations that favor them. Generally speaking, in presence of a short side chain, typically in the case of Phe-Ala or Phe-Gly sequences or the reverse ones, folded β-turn forms are able to compete with the extended structures (5-7 or 7-7).[26,29,31,42] In the presence of the Ser and Cys residues, the intra-residue interactions (Fig. 1) must be considered: *i.e.* OH/SH···OC_i (C6) and NH_i···O/S (C5). The examination of the interatomic distances for the most stable forms calculated (Table 3) enables us to rationalize the occurrence of the observed interactions (Fig. 4) as a function of the main chain folding on the Ser/Cys residue as well as the experimental effects of the energetic competition observed, namely the apparent destabilization of β-turns with Ser.

– Case of Ser: when Ser is in a γ-turn conformation (in FCa 5-7 or CFa 7-7 forms), short C6 intra-residue bonds can be formed, with a C6 distance of the order of 220 pm. When Ser is embedded in the central part of a β-turn as the second residue, the backbone conformation is much less favorable to these C6 interactions, which are then characterized by a much elongated C6 distance (>250 pm). Still worse, when in first position, the bond is inexistent. Besides, in all cases, a weak C5 interaction also takes place with relatively large C5 distances, between 240 and 265 pm, depending on the backbone conformation. From this, one can anticipate an extra stabilization of the 5-7 and 7-7 backbone structures at the expense of β-turns, in agreement with the calculated energies (Tables 1 and 2) together with the experimental observation: no β-turn could be experimentally detected with Ser (Fig. 1).

– With Cys, the larger size of the sulfur atom enables to form short C5 intra-residue interactions between NH(i) and S in all

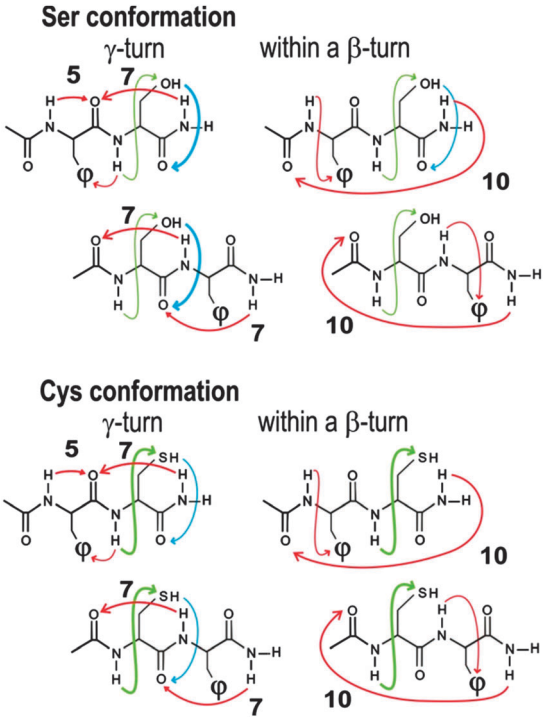

Fig. 4 Synthetic description of the relevant H-bonding schemes for the eight types of structures anticipated. Backbone H-bond are color-coded in red, whereas donor and acceptor side chain H-bond are depicted in blue and green respectively. The width of the arrows indicates the strength of the interactions. The lack of strong C5 or C6 interactions in the β-turns of Ser structures explains their destabilization, and their absence in the UV spectra.

Table 3 Structural parameters (RI-B97-D/TZVPP level of theory) of the double γ-turn and β-turn conformations in competition in the four sequences examined: conformations are identified from the terminology used in Tables 1 and 2, their Gibbs energy (kJ mol^{-1}), Ramachandran φ, ψ and side chain orientation χ_1 dihedral angles of the Ser/Cys residues are given in degrees, and H-bond distance for the intra-residue interaction with the CO(i) and NH(i) groups in pm. Values in bold correspond to observed conformations

			Serine-containing sequences						Cysteine-containing sequences				
	Conf.	ΔG (300 K)	φ	ψ	χ_1	OH–OC*i*	NH*i*–O_{Ser}	ΔG (300 K)	φ	ψ	χ_1	SH–OC*i*	NH*i*–S_{Cys}
γ-Turn FX	**5-7_1**	**0**	**−80**	**71**	**55**	**220**	**242**	**2.6**	**−80**	**67**	**50**	**242**	**263**
β-Turn FX	10_1	6.5	−107	25	64	252	265	**0**	**−102**	**13**	**59**	**288**	**277**
	10_2	11.1	−103	27	64	250	266	1.6	−94	11	59	292	278
	10_3	16.4	−115	14	46	—	251	5.6	−112	16	54	—	278
γ-Turn XF	**7-7_1**	**0**	**−82**	**67**	**56**	**218**	**246**	**2.4**	**−82**	**62**	**52**	**241**	**266**
β-Turn XF	**10_1**	5.7	−71	−10	56	329	248	**0**	**−70**	**−10**	**57**	**333**	**273**
	10_2	4.6	−66	−17	53	—	245	2.8	−70	−12	56	—	**272**
	10_3	10.9	−67	−18	−54	—	248	6.6	−67	−19	−56	—	274
	10_4	9.5	−71	−8	56	—	240	7.2	−71	−8	57	—	274

the conformations (5-7, 7-7 and β-turn), with C5 distances of the order of 260–280 pm despite a marked deviation from linearity (N–H–S angle ~110°). The SH–O C6 interactions, in contrast, have a much lesser strength, as testified by the long interatomic distances (>240 pm), much larger than in the Ser case. The prominent character of the C5 bonds together with the minor effect of C6 bonds does not generate any substantial differential stabilization between (5-7 or 7-7) extended forms and β-turns, leaving the latter energetically competitive, in agreement with both calculated stabilities (Table 3) and experimental detection of both β-turns and extended forms.

The balance between extended and folded forms results from the interplay between the intra-backbone interactions and the side-chain backbone intramolecular interactions. It appears to be mediated by the nature of the intra-residue interactions, more precisely by the balance between the C6 interaction favored with Ser, and the C5 interaction favored in Cys (*cf.* illustration in Fig. 4). The nature of the heteroatom controls the H-bond donor and acceptor properties of the Ser/Cys side chain: OH is a better donor than SH, but the opposite is true for the relative accepting characters of O and S. The present gas phase experiments, where the effects of individual interactions can be quantified, elegantly document this interplay in specific backbone conformations (γ- or β-turns). One can anticipate that this effect is probably more general and plays a role in stabilizing conformations, beyond those stabilized in the gas phase.

Spectroscopic probe of the intraresidue NH–O/S C5 H-bonds

The present experiments show that gas phase spectroscopy provides an overall picture of the intra-residue interactions and of their change when going from Ser to Cys. The C5 interaction, in particular, can be easily tracked, as illustrated in Fig. 5. The IR spectra (Fig. 3) are compared with those of conformations having the same backbone H-bonding structure in more simple capped dipeptides, *i.e.*, whose side chain does not introduce any additional interaction (Ala and Gly).[26,31,42] As a matter of fact, the strength of the C5 bond can be estimated from the spectral situation, where the corresponding NH bond does not experience any local perturbation (free NH stretch, in the 3490 cm^{-1} region). This is particularly obvious when Phe is in second position. When the residue is in a γ-turn conformation (Fig. 5, top panel), the red shift of the $NH_{Ser/Cys}$ band is of the order of 40 cm^{-1} with Ser and increases to 75 cm^{-1} with Cys, featuring a much stronger interaction. The same trend is observed when Phe is in first position (Fig. 5, lower panel) although in this case, the $NH_{Ser/Cys}$ experiences the influence of both the O/S lone pairs and the π phenyl ring.

Comparison of γ-turn and β-turn conformations for Cys (Fig. 5, top and central panel) is also enlightening: the red shift of the NH_{Cys} band of 75 cm^{-1} for the γ-turn drops to 50 cm^{-1} in the β-turn, in agreement with the C5 distances calculated to be 266 and 273 pm respectively (Table 3), *i.e.*, slightly longer than the optimum distance NH–S distance within the formamide–MeSH complex.[13] Concerning the spectral linewidths, these C5 interactions of Cys, similarly to NH–π interactions, lead to narrow lines despite a significant strength, suggesting a moderate coupling of the NH stretch mode with the low frequency modes of the backbone, in contrast for instance with the C7 strong H-bonds of γ-turns.[33]

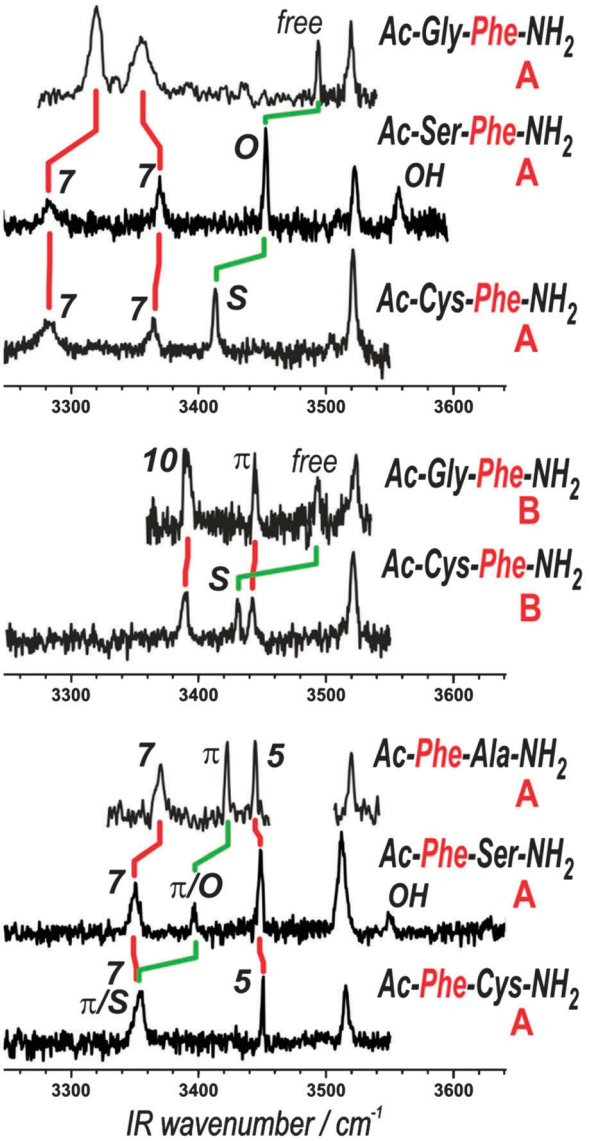

Fig. 5 Comparison of the IR/UV double resonance spectra of conformers having the same backbone H-bonding pattern in a series of FXa or XFa model molecules (top panel: 7-7; center: 10; bottom: 5-7). The evolution of the most relevant interactions along each series is indicated by colored lines (in red for 5, 7 or π backbone interactions, in green for the C5 interaction).

The C6 interaction of Ser can also be easily be tracked for Ser in a γ-turn conformation. Fig. 5 shows similar positions for the OH_{Ser} for the FSa and SFa A forms, significantly red-shifted compared to a free OH in agreement with calculated H-bond distances (Table 3).

This spectroscopic approach is complemented by a visualization of the interactions, as obtained by the NCI-plot procedure,[43–45] which emphasizes the regions of space where the reduced gradient of the density function is vanishing (Fig. 6).

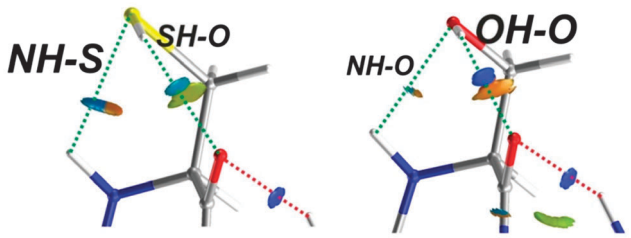

Fig. 6 Interactions between NH/OH/SH groups and accepting groups (O and S lone pairs) in the conformer 77-1 (A) of the Ac-Cys-Phe-NH_2 (left) and Ac-Ser-Phe-NH_2 (right) molecules, as revealed by colored isosurfaces (s = 0.35) of the electron density reduced gradient s (3D-NCI-plot; see text), following the NCI-plot topological analysis of the electron density at the RI-DFT-D/TZVP level of theory.

The type and strength of local interactions (as revealed by the critical electron density at this loci) are color-coded ranging from blue for strong stabilizing interactions to green for weak interactions; the red color being associated to destabilizing interactions, usually ring closures. The presence of a red/green surface accompanying the blue ones is due to a specific topological property of the density function induced by the closure of the intramolecular C5 rings.[45] The darker blue of the OH–O interaction indicates a higher critical electron density (0.0165 atomic units) compared to the SH–O interaction (0.0115 a.u.), showing the larger strength of the C6 interaction in Ser, in agreement with the fact that Ser is considered as better donor than acceptor. As for the C5 interaction, the isosurface found for Cys, at a critical density of 0.0135 a.u., indicates a significant interaction, in contrast with the negligibly small isosurface in Ser at a lower density (0.0115 a.u.). These values corroborate the ranking of the NH–S/O and SH/OH–O interactions in Cys and Ser as obtained from the IR spectroscopy. Again, this fits well to the better acceptor properties of Cys compared to Ser.

The C5 intraresidue interaction of Cys can be compared to a similar bonding pattern in related compounds, in particular the methionine residue. Thus the C6 intraresidue H-bond of methionine linking NH_{Met} to the Met side chain S sulphur atom has been characterised[28] by a larger red-shift (with a NH stretch frequency value in the 3360 cm^{-1} range), shorter calculated NH–S distances of $\sim$245 pm and more linear bonding geometries (N–H–S angle θ in the 130–140° range). Obviously, the additional methylene spacer the Met side chain, provides an additional degree of freedom, facilitating the formation of a stronger H-bond. By contrast, such Met reference species demonstrates the effects of constraints within the shorter Cys side chain.

Concerning the comparison with protein structures, it is interesting to bear in mind that most of the interactions ($\sim$3/4) of cysteine side chains in proteins are close contacts between N_i and S_i atoms of the same Cys residue, described by Zhou *et al.*[3] as ''small-angle interactions''. The calculated NH_i–S_i distances and χ_1 N_{Cys}–Cα–Cβ–S dihedral angles in the Cys residues forms encountered in the present experiment (Fig. 3 and Table 3: $\sim$270 pm, $\chi_1 \sim 55°$. see Table 3) are very close to those typically found for the ''small-angle interactions'' ($\sim$280 pm, and $\chi_1 \sim 60°$), thus providing a strong physical ground for their prevalent character in proteins.

Spectroscopic probe of the interresidue O/S–π interactions

The spectral shifts of the origin band of conformer A along the FAa, FSa and FCa series, shown in Fig. 2 also deserves discussion. The origin transition of the π system provides an indirect probe of the interactions undergone by this moiety. This transition being controlled by a differential effect between the ground and the first excited state of a ππ* nature, a red shift suggests that the interaction undergone does increase upon excitation. This is the case for instance in dispersive interactions: their magnitude, which goes with the polarisability of the moieties involved, is larger in the excited state because of the larger polarisability induced by a large diffuseness. For example a nearly purely dispersive interaction, such as toluene–rare gas, gives rise to red shifts of the order of a few tens

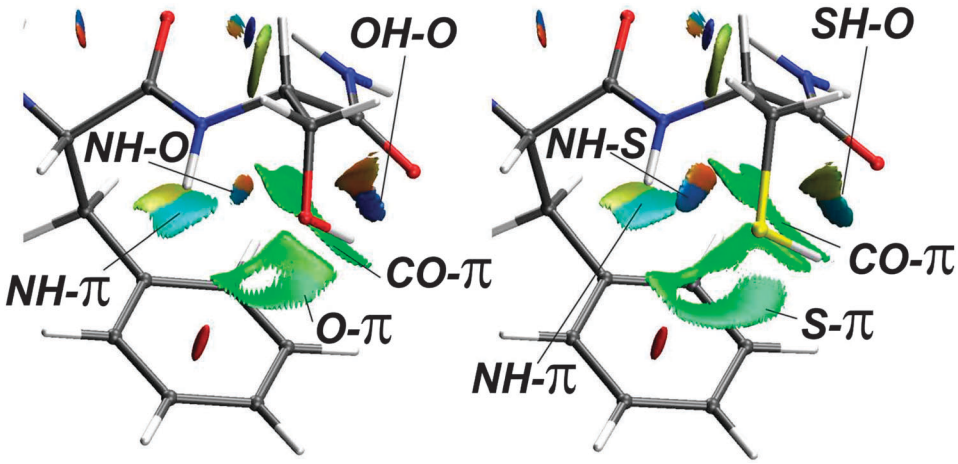

Fig. 7 Interactions involving the phenyl π ring, the NH/OH/SH groups and the side chain accepting sites (O and S lone pairs) in the conformer 5-7_1 (A) of the Ac-Phe-Ser-NH_2 (left) and Ac-Phe-Cys-NH_2 (right) molecules, as revealed by colored isosurfaces (s = 0.35) of the electron density reduced gradients (3D-NCI-plot; see text), following the NCI-plot topological analysis of the electron density at the RI-DFT-D/TZVP level of theory.

of cm^{-1} (25 for Ar, 39 for Kr).[46] Red-shifts comparable to those of Fig. 2 were also observed in FXa dipeptide systems having the same 5-7 structure, but for which the X side chain was bulky enough to cause close contacts.[29,33] In the present case the interactions between the lone pair of the O/S side chain heteroatom and the π cloud presumably obey such a rule. In other words, we suggest that the conformer A red shifts (of 66 and 132 cm^{-1} for Ser and Cys respectively) in Fig. 2 can be assigned to interresidue O/S–π interactions. This view is corroborated by the NCI analysis of these secondary interactions involving the phenyl ring (Fig. 7). Apart from the NH–π interaction evidenced by a bicolored isosurface,[45] two extended green sheets, corresponding to the close contacts of the π system with the O/S heteroatoms as well as the nearby Ser/Cys carbonyl, are found on the 3D-NCI plot. These surfaces are typical of dispersive interactions, as exemplified in the benzene or methane dimers.[44,45] The strength of the NH–π interaction being similar in Phe–Ser and Phe–Cys (as illustrated by the same color of the corresponding isosurface in Fig. 7), the larger sheet with Cys should be ascribed to a larger dispersive interaction with the more polarizable S atom. This present finding of close contacts between electron-rich moieties (O/S lone pair *vs.* π ring) echoes similar situations encountered in proteins, as revealed from data mining on protein data with cystine (S–S bridges) or methionine residues,[2] both containing S lone pairs similar to Cys in the close vicinity of an aromatic ring.

Conclusion

Gas phase peptide chains are shown to be useful models to precisely document local interactions taking place in proteins. The local nature of these interactions *de facto* validates such an isolated molecule approach, which also enables the acute insight of laser spectroscopy, both in the IR and UV spectra ranges. The resolution achieved in cold peptide chain enables spectroscopists to track the individual spectral features of each NH stretch in the molecule. Then the quantum chemistry calculations at hand for these systems and the scaling procedures developed are mature enough to enable us to propose a structural assignment using an automated procedure, corroborated by an alternative approach based on comparison with similar systems.

In these model systems the presence of OH/SH–OC(i) C6 and NH(i)–O/S C5 intraresidue interactions in Ser/Cys is demonstrated from laser spectroscopic probing. The strength of these local interactions linking side chain and main chain, is documented from their IR spectral features for well-defined conformations of the main chain (Cys and Ser, in either a γ-turn or a β-turn). The spectroscopic analysis demonstrates that a subtle competition exists between the two types of intra-residue bond: the C6 H-bond is the major interaction with Ser, in contrast to Cys where C5 interaction takes over; their strength being dependent on the residue (Cys *vs.* Ser) as well as on the backbone conformation. As a more general outcome, the study emphasizes the role of these ancillary *local* interactions, which do participate in the overall conformational backbone stability in peptides and proteins and significantly mediate the competition between main chain folding patterns, *e.g.*, folded *vs.* extended forms. It also provides chemical physics grounds to the specific structures found in X-ray diffraction protein structure analysis like the N–S close contacts evoked by Zhou *et al.* in their study[3] of Cys residues in proteins. Clearly these local C6 and C5 bonding patterns are local motifs to be considered when modelling or analyzing proteins' structure.

Acknowledgements

MA acknowledges financial support for his post-doctoral fellowship from the Fondation de Coopération Scientifique ''Campus Paris-Saclay'', through the Inter-Labex PALM-CHARMMMAT Grant ''LALIFOLD2013-0272I'' under the Program *Investing in the Future* (Grant no. ANR-11-IDEX-0003-02), funded by the French National Research Agency.

Notes and references

1 P. Chakrabarti and D. Pal, *Prog. Biophys. Mol. Biol.*, 2001, **76**, 1–102.
2 P. Chakrabarti and R. Bhattacharyya, *Prog. Biophys. Mol. Biol.*, 2007, **95**, 83–137.
3 P. Zhou, F. Tian, F. Lv and Z. Shang, *Proteins: Struct., Funct., Bioinf.*, 2009, **76**, 151–163.
4 N. Eswar and C. Ramakrishnan, *Protein Eng.*, 2000, **13**, 227–238.
5 D. Pal and P. Chakrabarti, *J. Biomol. Struct. Dyn.*, 1998, **15**, 1059–1072.
6 E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2001, **3**, 1–18.
7 T. S. Zwier, *J. Phys. Chem. A*, 2006, **110**, 4133–4150.
8 J. P. Simons, *Mol. Phys.*, 2009, **107**, 2435–2458.
9 M. S. de Vries and P. Hobza, *Annu. Rev. Phys. Chem.*, Annual Reviews, Palo Alto, 2007, vol. 58, pp. 585–612.
10 T. N. Wassermann, D. Luckhaus, S. Coussan and M. A. Suhm, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2344–2348.
11 D. L. Howard and H. G. Kjaergaard, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4113–4118.
12 P. Du, X. K. Jiang and Z.-T. Li, *Tetrahedron Lett.*, 2009, **50**, 320–324.
13 B. J. Mintz and J. M. Parks, *J. Phys. Chem. A*, 2012, **116**, 1086–1092.
14 V. S. Minkov and E. V. Boldyreva, *J. Phys. Chem. B*, 2013, **117**, 14247–14260.
15 V. S. Minkov and E. V. Boldyreva, *J. Phys. Chem. B*, 2014, **118**, 8513–8523.
16 E. A. Kapustin, V. S. Minkov, J. Stare and E. V. Boldyreva, *Cryst. Growth Des.*, 2014, **14**, 1851–1864.
17 H. S. Biswal, S. Chakraborty and S. Wategaonkar, *J. Chem. Phys.*, 2008, **129**, 184311.
18 H. S. Biswal and S. Wategaonkar, *J. Phys. Chem. A*, 2009, **113**, 12763–12773.

19 H. S. Biswal and S. Wategaonkar, *J. Phys. Chem. A*, 2009, **113**, 12774–12782.
20 H. S. Biswal, P. R. Shirhatti and S. Wategaonkar, *J. Phys. Chem. A*, 2009, **113**, 5633–5643.
21 H. S. Biswal and S. Wategaonkar, *J. Phys. Chem. A*, 2010, **114**, 5947–5957.
22 H. S. Biswal and S. Wategaonkar, *J. Chem. Phys.*, 2011, **135**, 134306.
23 N. O. B. Luettschwager, T. N. Wassermann, S. Coussan and M. A. Suhm, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8201–8207.
24 A. Chakraborty, N. Guchhait, K. Le Barbu-Debus, A. Mahjoub, V. Lepere and A. Zehnacker-Rentien, *J. Phys. Chem. A*, 2011, **115**, 9354–9364.
25 A. Mahjoub, K. Le Barbu-Debus and A. Zehnacker, *J. Phys. Chem. A*, 2013, **117**, 2952–2960.
26 W. Chin, F. Piuzzi, I. Dimicoli and M. Mons, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1033–1048.
27 M. Gerhards, in *Principles of Mass Spectrometry Applied to Biomolecules*, ed. J. Laskin and C. Lifshitz, Wiley & Sons, Hoboken, NJ, 2006, pp. 3–62.
28 H. S. Biswal, E. Gloaguen, Y. Loquais, B. Tardivel and M. Mons, *J. Phys. Chem. Lett.*, 2012, **3**, 755–759.
29 E. Gloaguen and M. Mons, *Top. Curr. Chem.*, DOI: 10.1007/128_2014_580.
30 B. Yan, S. Jaeqx, W. J. van der Zande and A. M. Rijs, *Phys. Chem. Chem. Phys.*, 2014, **16**, 10770–10778.
31 W. Chin, J. P. Dognon, F. Piuzzi, B. Tardivel, I. Dimicoli and M. Mons, *J. Am. Chem. Soc.*, 2005, **127**, 707–712.
32 W. Chin, M. Mons, J.-P. Dognon, F. Piuzzi, B. Tardivel and I. Dimicoli, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2700–2709.
33 E. Gloaguen, F. Pagliarulo, V. Brenner, W. Chin, F. Piuzzi, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2007, **9**, 4491–4497.
34 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.
35 H. S. Biswal, Y. Loquais, B. Tardivel, E. Gloaguen and M. Mons, *J. Am. Chem. Soc.*, 2011, **133**, 3931–3942.
36 E. Gloaguen, Y. Loquais, J. A. Thomas, D. W. Pratt and M. Mons, *J. Phys. Chem. B*, 2013, **117**, 4945–4955.
37 E. Gloaguen, R. Pollet, F. Piuzzi, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2009, **11**, 11385–11388.
38 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
39 *HyperChem, version 7.52*, 2005, Hypercube.
40 *TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007*, available from http://www.turbomole.com, 2011.
41 W. Chin, F. Piuzzi, J.-P. Dognon, I. Dimicoli and M. Mons, *J. Chem. Phys.*, 2005, **123**, 084301.
42 W. Chin, J.-P. Dognon, C. Canuel, F. Piuzzi, I. Dimicoli, M. Mons, I. Compagnon, G. von Helden and G. Meijer, *J. Chem. Phys.*, 2005, **122**, 054317.
43 E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen and W. Yang, *J. Am. Chem. Soc.*, 2010, **132**, 6498–6506.
44 J. Contreras-Garcia, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, 2011, **7**, 625–632.
45 R. Chaudret, B. de Courcy, J. Contreras-Garcia, E. Gloaguen, A. Zehnacker-Rentien, M. Mons and J. P. Piquemal, *Phys. Chem. Chem. Phys.*, 2014, **16**, 9876–9891.
46 M. Mons, J. Le Calvé, F. Piuzzi and I. Dimicoli, *J. Chem. Phys.*, 1990, **92**, 2155.

