PCCP

PAPER



Cite this: DOI: 10.1039/c6cp04909e

Received 14th July 2016, Accepted 25th August 2016

DOI: 10.1039/c6cp04909e

www.rsc.org/pccp

I. Introduction

The intrinsic conformational propensities of amino acid residues for promoting (or disrupting) helical secondary structure have been examined computationally^{1,2} and experimentally.³⁻⁶ The great majority of proteinaceous helices determined in X-ray crystal structures are right-handed α -helices,^{7,8} characterized by 3.6 residues per turn and hydrogen-bonded rings of 13 atoms (3.6₁₃-helices) closed by backbone hydrogen-bonding pairs between residues $i + 4 \rightarrow i$ and prototypical Ramachandran angles of ϕ , $\psi = -60^{\circ}$, -45° .⁷ About 10% of helices found in protein structures are more tightly-wound 3₁₀-helices, with 3.0 residues per turn and rings of 10 atoms closed by $i + 3 \rightarrow i$ backbone hydrogen-bonding partners and Ramachandran angles of ϕ , $\psi = -57^{\circ}$, -30° .⁸⁻¹⁰ Although the α -helix is the



Joseph R. Gord,^a Daniel M. Hewett,^a Alicia O. Hernandez-Castillo,^a Karl N. Blodgett,^a Matthew C. Rotondaro,^b Adalgisa Varuolo,^b Matthew A. Kubasik*^b and Timothy S. Zwier*^a

The conformational preferences of a series of capped peptides containing the helicogenic amino acid aminoisobutyric acid (Aib) (Z-Aib-OH, Z-(Aib)₂-OMe, and Z-(Aib)₄-OMe) are studied in the gas phase under expansion-cooled conditions. Aib oligomers are known to form 3_{10} -helical secondary structures in solution and in the solid phase. However, in the gas phase, accumulation of a macrodipole as the helix grows could inhibit helix stabilization. Implementing single-conformation IR spectroscopy in the NH stretch region, Z-Aib-OH and Z-(Aib)₂-OMe are both observed to have minor conformations that exhibit dihedral angles consistent with the 3_{10} -helical portion of the Ramachandran map (ϕ , $\psi = -57^{\circ}$, -30°), even though they lack sufficient backbone length to form 10-membered rings which are a hallmark of the developed 310-helix. For Z-(Aib)4-OMe three conformers are observed in the gas phase. Single-conformation infrared spectroscopy in both the NH stretch (Amide A) and C=O stretch (Amide I) regions identifies the main conformer as an incipient 310-helix, having two free NH groups and two C10 H-bonded NH groups, labeled an F-F-10-10 structure, with a calculated dipole moment of 13.7 D. A second minor conformer has an infrared spectrum characteristic of an F-F-10-7 structure in which the third and fourth Aib residues have ϕ , ψ = 75°, -74° and -52°, 143°, Ramachandran angles which fall outside of the typical range for 3_{10} -helices, and a dipole moment that shrinks to 5.4 D. These results show Aib to be a 3_{10} -helix former in the gas phase at the earliest stages of oligomer growth.

> most commonly found secondary structure, it accounts for only about 30% of residues in protein X-ray structures.⁸ Clearly, the forces that favor the development of well-defined helical structures, such as intramolecular hydrogen bonding and interresidue side-chain packing, are competing against other conformational imperatives. In fact, the balance between these forces can sometimes be tipped by the presence of even a single residue.^{11,12}

> Aminoisobutyric acid (Aib) is a naturally occurring amino acid containing two side chain methyl groups. The symmetric dimethylation of the α -carbon renders Aib achiral, similar to glycine. In stark contrast to glycine, which is known to be the residue least supportive of helical structure,^{1,3–6,10,13} Aib residues are strongly helicogenic; oligomers of Aib are known to induce formation of the 3₁₀-helix in solution and in the condensed phase.^{13–16} The Aib residue is found singly and contiguously (up to (Aib)₄) in the naturally occurring peptaibol family of antimicrobial peptides. The helicogenicity of contiguous Aib residues is believed to allow peptaibols to form helical transmembrane pores essential to their mechanism of antimicrobial activity.^{13,15,17}



View Article Online

^a Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA. E-mail: Zwier@purdue.edu

^b Department of Chemistry and Biochemistry, Fairfield University, Fairfield, Connecticut 06824, USA. E-mail: MKubasik@fairfield.edu

[†] Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp04909e

Paper

The Amide A and Amide I vibrational modes of polar amide functional groups are particularly sensitive to the local electrostatic environment and are consequently diagnostic of protein and peptide secondary structural features such as intramolecular hydrogen bonding.^{18,19} Toniolo and coworkers monitored the onset of helical secondary structure in a series of Aib homopeptides, Z-(Aib)_n-OtBu (Z = benzyloxycarbonyl cap, n = 1-12), by examining the NH stretch (Amide A) region of the infrared.¹⁴ Since the NH stretch fundamentals shift to lower frequency with increasing strength of the hydrogen bonds in which they are involved, they serve as an important diagnostic of structure, especially in non-polar solvents where interference from solvent is minimal.¹⁹ Solvated in deuterochloroform, Z-(Aib)_n-OtBu exhibits two amide A spectral bands, assigned to "free" and "intramolecularly hydrogen bound" NH stretches. With increasing 'n', the "hydrogen bound" spectral band grows at the expense of the "free" spectral band, converging to a limit of two free NH stretches, characteristic of a 310-helical bonding pattern.

In a more recent study, Toniolo and Ge employed 2D IR spectroscopy to investigate the C=O stretch (Amide I) spectroscopy of a series of Aib oligomers $(Z-(Aib)_n-OtBu, n = 3, 5, 8, 10)$ dissolved in CDCl₃.²⁰ Although individual modes are not resolvable in CDCl₃, estimates of local Amide I frequencies and coupling constants have been obtained through fits of the 2D IR spectral contours. The 2D IR spectra indicate that Amide I spectral features characteristic of 3_{10} -helices emerge as early as $n = 5.^{20}$ The isotope-edited Amide I region linear infrared spectra of Kubasik and coworkers, employing ¹³C-enriched Z-(Aib)₆-OtBu isotopologues, reveal localized ¹³C=O Amide I modes as distinct spectral features at reduced frequencies-an improvement over the broad, unresolved ¹²C=O Amide I band.²¹ These isotopeedited spectra confirm a 310-helical hydrogen-bonding pattern for Z-(Aib)₆-OtBu, since the lowest frequency ¹³C-localized Amide I oscillators are those that, within a 310 helical hydrogen-bonding pattern, are both hydrogen bond acceptors (at the carbonyl oxygen) and hydrogen bond donors (at the amide group hydrogen). Although ¹³C-enrichment shifts a ¹³C-Amide I band to lower frequency, the spectral features in the Amide I and Amide A regions are still inherently broadened by solvent effects.

Under jet-cooled conditions, gas phase spectroscopic studies of peptides offer superior spectral resolution as compared to solvent-broadened condensed phase efforts. Additionally, gas phase studies examine the intrinsic conformational propensities of amino acid residues and peptides by removing subtle structuresupporting aspects of a solvent environment. Furthermore, in the gas phase, the NH stretch region can be used as a structural diagnostic, reporting directly on the hydrogen-bonding architecture of the conformations, with no interference from solvent absorption (*e.g.*, due to H_2O).

In a recent study of the glycine homopeptide series Z-(Gly)_{*n*}-OH with n = 1, 3, 5 and Z-(Gly)₅-NHMe (where Z is a benzocarboxy cap, Fig. 1), Dean *et al.* showed that both of the n = 5 oligomers folded exclusively into 14/16 mixed helices in which the hydrogen bonds associated with 14-membered H-bonded rings (labeled C14) and 16-membered rings (C16) point in opposite directions along the peptide backbone. The assigned structures possess



Fig. 1 The complete series of capped Aib homopeptides investigated in this study. C=O groups have been numbered in red and NH groups have been numbered in blue, in both cases from N- to C-terminus with 0 indicating the C=O in the benzyloxycarbonyl cap (Z-cap).

calculated permanent dipole moments of only 2.6 and 2.8 D for the –OH and –NHMe structures, respectively,²² values smaller than the dipole moments of the individual peptide units (3.46 D). The conformational flexibility of the glycine residue allowed the backbone to adopt a configuration capable of supporting this mixedhelical architecture, with C7 ($i + 2 \rightarrow i$) H-bonds serving as spacers between the remote C14 and C16 H-bonds.

Gas phase studies have also been carried out to evaluate the competition among different hydrogen-bonding motifs in a series of tetrapeptides capable of forming a variety of intramolecular hydrogen bonds.²³ Mons and coauthors saw competition among structures involving C10 and C7 hydrogen-bonded rings, $NH \cdots \pi$ hydrogen-bonded interactions, and 2₇-ribbon structures in their study of Ac–Phe–Ala–Ala–NH₂, Ac–Ala–Ala–Phe–NH₂, and Ac–Ala–Phe–Ala–NH₂.

Additionally, Mons and co-workers have examined a tetrapeptide containing a phenylalanine residue flanked by two Aib residues (Ac–Aib–Phe–Aib–NH₂) under jet-cooled conditions.²⁴ Its main conformer exhibited NH stretch transitions due to a pair of C10 H-bonds that are indicative of an incipient 3_{10} -helix, appearing at 3374 cm⁻¹ and 3405 cm⁻¹.^{22,24,25}

The goal of the present study is to investigate more fully the folding propensities and spectroscopic signatures of the series of capped Aib homopeptides shown in Fig. 1, under cold, isolated conditions in the gas phase. These structures will be referred to in short-hand as Z-(Aib)_n-X (n = 1, 2, 4 and X = OH [for n = 1 only] or OMe). The benzyloxycarbonyl (Z) cap provides the necessary UV chromophore for carrying out single-conformation spectroscopy, while retaining a pure Aib homopeptide backbone. The present work examines the intrinsic conformational propensities of the Aib residue and contiguous Aib units, elaborated as short homooligomers. By cooling the molecule in a supersonic expansion, the present work probes the lowest energy conformations of these homopeptides in their isolated form, unperturbed by solvent or solid state intermolecular interactions. Single-conformation IR spectra are compared to the predictions of DFT calculations, leading to assignments that identify the preferred conformations of the molecules in this series, and spectroscopically characterize them in the ultraviolet (UV) and infrared (IR).

II. Methods

A. Experimental methods

Samples were laser desorbed into the gas phase from the surface of a graphite rod using the 1064 nm output from a Nd:YAG laser (Continuum MiniLite).²⁶ Laser pulses were approximately 6 ns in duration with 3-5 mJ of energy per pulse. The desorption rod was translated linearly using a pusher motor during data collection to ensure that fresh sample was available throughout each scan. This laser desorption scheme has been described in detail elsewhere.^{22,27} After being desorbed, the gas-phase sample molecules were entrained in an expansion of argon gas coming from a Parker Fluidics, Series 9 general valve with a 500-800 µm nozzle diameter operating at 20 Hz with approximately 5.5 bar of backing pressure. Sample molecules were then collisionally cooled in a supersonic expansion, funneling their population primarily into the zero-point vibrational energy levels of the most stable conformers.^{28–30} Downstream, the seeded expansion was passed through a 3 mm skimmer to form a molecular beam which was interrogated using single- and double-resonance laser schemes in the ionization region of a Wiley McLaren, time-of-flight mass spectrometer.31

Resonant two-photon ionization (R2PI) was used to record the electronic excitation spectra for the molecules of interest. Scans over the 35 000 cm⁻¹ to 38 000 cm⁻¹ region were recorded using the doubled output of a Radiant Dyes, NarrowScan dye laser pumped by 355 nm light from a Continuum, SureLite II, Nd:YAG laser. When UV light from the dye laser is resonant with a vibronic transition between the S_0 and S_1 states of a given conformation, a first photon is absorbed which promotes the molecule to the excited state; upon immediate absorption of a second photon, the molecule is ionized. By tuning the UV laser over a range of wavelengths characteristic of the Z-cap UV chromophore, the vibronic spectrum of all conformations present in the jet can be recorded.

A KTP/KTA-based infrared parametric converter (LaserVision) was used to generate IR light in the NH and OH stretch regions (3200 cm⁻¹ to 3700 cm⁻¹) with typical output energies on the order of 12–15 mJ per pulse. When desired, a AgGaSe₂ crystal was used to difference frequency mix into the mid-infrared, in order to generate light in the 1600 cm⁻¹ to 1800 cm⁻¹ region which encompasses the Amide I (C=O stretch) fundamentals.³² Typical output energies in the Amide I region were on the order of 500–1000 μ J per pulse.

Single-conformation IR spectra were recorded using resonant ion-dip IR spectroscopy (RIDIRS).^{32,33} In RIDIRS, the UV laser, operating at 20 Hz, has its wavelength fixed on a vibronic transition belonging to a single conformer, selected from the R2PI spectrum, in order to generate a constant level of ion signal. The 10 Hz, infrared beam is spatially overlapped with, but temporally precedes the UV beam. When the IR laser is resonant with a vibrational transition belonging to the same conformation that is being pumped with the UV laser, some of the population is moved out of the vibrational zero-point level and promoted to a vibrationally excited level within the ground electronic state, thus the total number of molecules available for the UV laser to ionize is decreased, and a dip in the total ion signal is observed. By performing active baseline subtraction on each pair of "IR on" and "IR off" peaks with a Stanford Research Systems gated integrator (SR-250), a spectrum can be recorded in which the fractional depletion in ion signal is plotted as a function of IR frequency.

Once the R2PI and RIDIR spectra have been recorded, IR-UV hole-burning (HB) spectroscopy can be used to separate the R2PI spectrum into a set of conformer-specific UV spectra.³² To collect these IR-UV HB spectra, the wavelength of the IR laser is fixed on a vibrational transition belonging to a specific conformation, and the wavelength of the UV laser is scanned. By implementing the same timing and overlap schemes described in RIDIRS, and performing active baseline subtraction, the fractional depletion of the ion signal *versus* UV frequency can be recorded to obtain single-conformation UV spectra.

B. Computational methods

The results of Density Functional Theory (DFT) calculations are used to assist in making structural assignments of the observed IR spectra. To obtain these computational results, a starting structure is submitted to a conformational search performed at the Amber* force field³⁴ level within a 100 kJ mol⁻¹ energy window using MacroModel (Schrödinger, Inc.).34,35 Structures from the conformational search with energies within ~ 50 kJ mol⁻¹ of the force field global minimum are re-optimized to calculate geometries and harmonic frequencies and IR intensities via Density Functional Theory (DFT) calculations (Gaussian09³⁶) using the M05-2X functional and a 6-31+G(d) basis set.³⁷ The vibrational frequencies are scaled using well established scale factors of 0.940 in the NH and OH stretching regions and 0.960 in the Amide I region for this level of theory.^{22,38-40} We would like to point out that there is also an opportunity to apply a slight correction ($\sim 10 \text{ cm}^{-1}$) to the amide NH stretch frequencies to account for stretch-bend Fermi resonance; however we have not done so here.⁴¹ Relative energies of the conformers, including corrections for zero-point energy effects at the harmonic level, are included in the figures and tables as appropriate.

C. Synthetic methods

Commercially available starting materials Aib (Acros), Z-Aib-OH (Bachem) and HCl·Aib-OtBu (Bachem) were used as received. HCl·Aib-OMe was prepared from MeOH and trimethylsilyl chloride.⁴² Aib dimers, Z-(Aib)₂-OtBu and Z-(Aib)₂-OMe, were prepared using EDCI as a coupling reagent at 60 °C in a CEM Discover SP microwave reactor for 30 minutes.⁴³ These two dimers were used to prepare Z-(Aib)₄-OMe *via* literature methods.¹⁴ Z-(Aib)₂-OMe and Z-(Aib)₄-OMe were purified *via* flash silica gel chromatography.

D. Nomenclature

Analysis of the single-conformation data employs the amide group numbering scheme shown in Fig. 1. In this scheme, NH groups are labeled 1–4 and C=O groups are numbered from 0–4 beginning at the N-terminal carbamate group of the Z-cap and extending to the C-terminus. Hydrogen-bonded rings are

named according to the number of atoms in the cycle closed by the H-bonding interaction. For example, a ten-atom cycle formed by an intramolecular hydrogen bond is labeled as 'C10'. To describe the hydrogen-bonding architecture of the conformations, a notation analogous to that used by Mons and co-workers⁴⁴ is implemented in which the hydrogen-bonding behavior of NH groups will be represented from N- to C-terminus by either "F" meaning a free NH, uninvolved in an intramolecular hydrogen bond, or "n", where "n" is the number of atoms in the resulting hydrogen-bonded ring. So, for example, a structure having 4 NH groups might be described as F-F-10-10 meaning the NH groups from N- to C-terminus are free, free, C10, and C10. It should be noted that, due to the achiral nature of the Aib residue, the structures shown in the figures and referred to in the tables have mirror images that are identical in energy and spectral features. The mirror image structures have backbone angles of opposite sign to those reported in the table(s).

III. Results

A. Z-Aib-OH

Resonant two-photon ionization (R2PI) and IR-UV hole-burning spectra for Z-Aib-OH are shown in Fig. 2(a). The R2PI spectrum contains a single dominant feature at 37 602 cm⁻¹ and two minor features at 37 576 cm⁻¹ and 37 619 cm⁻¹. The IR-UV HB spectra identify the 37 602 cm⁻¹ and 37 576 cm⁻¹ transitions as S_0 - S_1 origin transitions for two distinct conformations of Z-Aib-OH, labeled A and B, respectively. The peak at 37 619 cm⁻¹ is identified as a vibronic band, tentatively assigned to a butterfly motion of conformer A, involving the ring and backbone, based upon our IR assignments discussed below. The origin of conformer A is shifted just 1 cm⁻¹ from the origin transition for Z-Gly-OH studied by Dean *et al.* (37 601 cm⁻¹). Additionally, the assigned vibronic activity is nearly identical to that for Z-Gly-OH suggesting that the two structures are very similar, with nearly equivalent environments for the Z-cap.²²

Resonant ion-dip infrared spectroscopy was used to record conformer-specific IR spectra in the NH and OH stretch regions for both conformers of Z-Aib-OH. Assignments are made by comparing experimental spectra with calculated, scaled, harmonic frequencies and IR intensities, depicted as stick spectra, in Fig. 2(b). The RIDIR spectrum for conformer A shows two wellresolved transitions in the NH/OH stretching region. Comparison to computational results allows us to assign the transition at 3447 cm⁻¹ to a C5 hydrogen-bonded NH group. The second transition, located at 3579 cm⁻¹, is assigned to a free OH stretch in the carboxylic acid group. The RIDIR spectrum for the minor conformer B is unusual in that it has two broadened transitions centered near those of conformer A as well as a pair of sharp transitions at 3468 cm⁻¹ and 3583 cm⁻¹ (marked with daggers in Fig. 2(b)). It is likely that the frequency of the UV monitor transition has contributions both from conformer B (responsible for the sharp IR transitions) and from a hot band of conformer A (responsible for the broadened bands centered on the transitions due to conformer A). The small shift and broadening in frequency is consistent with excitation out of a vibrationally excited level of a low-frequency vibration, which can shift and modulate the NH stretch frequency. The small depletions due to B are also consistent with this overlap in the UV, where a portion of the ion signal is due to 'A', and therefore not available for depletion on resonances due to 'B'.



Fig. 2 (a) R2PI (black) and IR-UV hole-burning spectra (red for A and blue for B) for Z-Aib-OH. Hole-burning spectra were recorded with the IR hole-burn laser fixed at 3579 cm^{-1} for A, and 3468 cm^{-1} for B. (b) RIDIR spectra for conformers A and B of Z-Aib-OH in the NH and OH stretch region. Scaled, harmonic, vibrational frequencies and infrared intensities for the assigned structures are shown as stick spectra below the experimental data, calculated at the DFT M05-2X/6-31+G(d) level of theory. The daggers indicate transitions specific to conformer B. Asterisks indicate frequencies used for IR-UV hole-burning experiments. See text for further discussion.

The assigned structures for conformers A and B are shown in Fig. 3, and key structural parameters are summarized in Table 1. Conformer A is calculated to be the global minimum, among all computed structures, and contains an extended C5 interaction between the NH and C=O groups that span the Aib residue with ϕ , $\psi = 180^{\circ}$, 180° and a *cis* carboxylic acid OH group. Conformer B is assigned to a significantly different structure in which neither of the C=O groups nor the NH group engage in hydrogen bonding. Instead, the steric crowding of the dimethylated α -carbon causes the two amide groups to orient with ϕ , $\psi = -49^{\circ}$, -37° , values close to those found in a 3_{10} -helix.¹⁰ Additionally, parallel alignment of the C=O groups produces a dipole moment that is twice the magnitude observed for conformer A (3.3 D in conformer B *vs.* 1.5 D in conformer A). Since neither conformer contains an interaction between the



Fig. 3 (a) Assigned structure for conformer A of Z-Aib-OH. The dotted red line indicates a C5 hydrogen-bonding interaction. (b) Assigned structure for conformer B of Z-Aib-OH. Structural families, computed (M05-2X/6-31+G(d)) relative energies (kJ mol⁻¹), and dipole moments (D) for the two conformers are listed.

B. Z-(Aib)₂-OMe

The R2PI and IR-UV HB spectra for Z-(Aib)₂-OMe are shown in Fig. 4(a). Most of the intensity in the R2PI spectrum appears below 37 500 cm⁻¹, significantly lower in energy than either of the origins for the two conformers of Z-Aib-OH (37602 and 37 576 cm⁻¹). IR-UV hole-burning spectra identify three conformations labeled A, B, and C shown in red, blue, and green, respectively. The S₀-S₁ origin of conformer A is assigned to the peak at 37 436 cm⁻¹. The hole-burning spectra for conformers B and C were a challenge to record due to the small intensity of their transitions relative to those of conformer A, requiring significant signal averaging to achieve an acceptable signal-tonoise ratio. Therefore, it is difficult to identify the S_0 - S_1 origin of conformer B with certainty, and the transition at 37545 cm^{-1} , the first transition relatively free of interference that shows up clearly in the R2PI spectrum, was used for RIDIRS. The S0-S1 origin of conformer C at $37\,604$ cm⁻¹, is within 2 cm⁻¹ of the origin for conformer A of Z-Aib-OH, indicating that the phenyl ring is likely located in an environment similar to the phenyl ring environment in conformer A of Z-Aib-OH.

The R2PI spectrum for Z-(Aib)₂-OMe is much more congested than that of Z-Aib-OH, due to the presence of several conformers and/or much stronger Franck–Condon activity in low-frequency modes. Specifically, a clear progression of peaks is seen in conformer A, beginning at 37 436 cm⁻¹, with members separated by approximately 13 cm⁻¹. Vibrational assignments, discussed below, indicate that this progression is caused by an in-phase rocking motion involving the phenyl ring and the C-terminal end of the backbone.

Fig. 4(b) presents RIDIR spectra of the three conformers of Z-(Aib)₂-OMe, recorded over the 3300–3500 cm⁻¹ region where the amide NH stretch fundamentals are known to appear. Methoxy substitution removes the carboxylic acid OH, so scans above 3500 cm⁻¹ were not necessary. The three IR spectra show distinct patterns characteristic of three structural types. Calculated (M05-2X/6-31+G(d)), scaled (0.940), harmonic frequencies and IR intensities are shown below the experimental data as stick spectra.

Table 1 Dihedral angles for all assigned conformations. ^a All angles are given in degrees. Relative energies are given in kJ mol ⁻¹											
Molecule	Relative energy (kJ mol ⁻¹)	H-bonding pattern	Ring orientation	ϕ (1)	ψ (1)	φ (2)	ψ (2)	φ (3)	$\psi(3)$	φ (4)	ψ (4)
Z-Aib-OH (A)	0	5	83	-180	-180						
Z-Aib-OH (B)	4.5	F-F	84	49	37						
Z-(Aib) ₂ -OMe (A)	0	F-5	71	-57	-38	-178	-179				
$Z-(Aib)_2$ -OMe (B)	7.2	F-7/5	77	-72	70	180	179				
$Z-(Aib)_2$ -OMe (C)	5.2	F-F	84	180	-180	180	180				
Z-(Aib) ₄ -OMe (S)	0.0	F-F-10-10	74	-61	-26	-55	-27	-56	-31	51	40
Z-(Aib) ₄ -OMe (NS)	3.5	F-F-10-10	74	-61	-27	-57	-22	-57	-35	-48	-40
Z-(Aib) ₄ -OMe (B)	3.9	F-F-10-7	69	-63	-25	-58	-30	75	-74	-52	143
Z-(Aib) ₄ -OMe (C)	24.4	14-7-7-F	92	-82	34	68	-79	-53	151	43	49

^a Due to the achiral nature of Aib, a second set of equivalent conformations exists in which the signs of all the dihedral angles are changed.



Fig. 4 (a) R2PI (black) and IR-UV hole-burning spectra for conformers A, B, and C of Z-(Aib)₂-OMe. The hole-burning spectra were recorded with the IR hole-burning laser fixed at 3418 cm⁻¹ for A, 3342 cm⁻¹ for B, and 3407 cm⁻¹ for C, marked with asterisks in (b). (b) RIDIR spectra for conformers A, B, and C of Z-(Aib)₂-OMe in the NH stretch region. Scaled, harmonic, vibrational frequencies and infrared intensities for the assigned structures calculated at the M05-2X/6-31+G(d) level of theory are shown as stick spectra below experiment. Transitions used to record RIDIR spectra are marked with asterisks in (a).

The RIDIR spectrum of conformer A shows two well-resolved NH stretch fundamentals at 3418 cm⁻¹ and 3472 cm⁻¹. The higher-frequency transition at 3472 cm⁻¹ is identified as the free stretch of the NH of the carbamate group resulting from the Z-cap. The lower-frequency peak at 3418 cm⁻¹ is assigned to a C5 hydrogen-bonded NH stretch fundamental of the interior amide group. This peak is shifted significantly (-29 cm^{-1}) from the position of the assigned C5 NH stretch in the IR spectrum of Z-Aib-OH conformer A. The C5 structure of Z-Aib-OH contains an NH from a carbamate group, while the C5 structure of Z-(Aib)₂-OMe, conformer A, contains an NH of an amide group. Conformer A is thus assigned to the F-5 structure shown in Fig. 5(a).

In the IR spectrum of conformer B, small transitions marked with daggers in Fig. 4(b) appear at frequencies due to the two

transitions of A, indicating a small overlap with a vibronic transition of conformer A at the UV wavelength chosen. Unique IR transitions at 3343 cm⁻¹ and 3461 cm⁻¹ are due to conformer B. The higher-frequency transition at 3461 cm⁻¹ is assigned to the free (N-terminal) NH group, while the lower-frequency 3343 cm⁻¹ peak is assigned to the central amide NH group that engages simultaneously in C7 and C5 hydrogen bonds with the Z-cap (carbamate) carbonyl and the C-terminal methyl ester carbonyl, respectively. The frequency of this NH stretch is close to that of other pure C7 hydrogen-bonded rings.^{22,45} The stick diagram shows the calculated, scaled spectrum that best matches the experimental spectrum, arising from an F-7/5, bifurcated double-ring structure, shown in Fig. 5(b).

The IR spectrum for conformer C contains two clearly resolved NH stretch fundamentals appearing at 3407 cm^{-1} and 3427 cm^{-1} .



Fig. 5 (a), (b), and (c) show the assigned structures for conformers A, B, and C of Z-(Aib)₂-OMe. Hydrogen bonds are indicated with dotted red lines. Structural families, relative energies (kJ mol⁻¹), and dipole moments (D) are listed.

The computed structure that offers the best match to this experimental spectrum exhibits two C5 hydrogen-bonded NH groups. As previously observed, the phenyl ring adopts an orientation similar to the extended structure (conformer A) of Z-Aib-OH, consistent with the electronic origins of these two structures being nearly identical. This structure is assigned as a fully-extended 5-5 structure. The N-terminal (carbamate) NH group is assigned to the lower-frequency transition at 3407 cm⁻¹, and the central amide NH group is assigned to the higher-frequency transition at 3427 cm⁻¹. This fully extended structure minimizes interaction between the side chains, backbone, and phenyl ring and has a dipole moment of 4.6 D. This is in line with the dipole moments calculated by Dean *et al.* for analogous extended structures of Z-Gly_n-OH.²²

The assigned structures for conformers A–C, shown in Fig. 5, possess the ϕ and ψ dihedral angles summarized in Table 1. It is noteworthy that the F-5 structure of conformer A, which is the global minimum structure, possesses Ramachandran angles $(\phi, \psi = -57^{\circ}, -38^{\circ})$ about Aib[1] that are near those found in an idealized 3_{10} -helix ($\phi, \psi = -57^{\circ}, -30^{\circ}$). The Ramachandran angles of Aib[2] allow for a C5 hydrogen-bonded structure. This same C-terminal C5 structure is found in the other two conformers, most likely because the C-terminal methyl ester possesses no NH group to participate in hydrogen bonding. The only other unique Ramachandran angles are those associated with the C7 γ -turn ($\phi_1, \psi_1 = +72^{\circ}, -70^{\circ}$), which despite its strong H-bond ($\nu_{\rm NH} = 3342 \text{ cm}^{-1}$) has an energy +7.2 kJ mol⁻¹ above the F-5 structure.

The significantly shifted electronic origins of conformers A and B as well as the extensive low-frequency vibronic activity observed in their R2PI spectra lend additional support to the conformational assignments in Fig. 5. Both structures fold the C-terminal end of the peptide back over the Z-cap, enabling a strong interaction so that upon π - π * excitation, Franck-Condon activity is evident, likely involving the low-frequency motions of the peptide backbone, ($\nu \sim 13 \text{ cm}^{-1}$) against the aromatic ring.

C. Z-(Aib)₄-OMe

Fig. 6 presents the R2PI and IR-UV hole-burning spectra of Z-(Aib)₄-OMe. The R2PI spectrum for Z-(Aib)₄-OMe shows significantly less congestion than that of Z-(Aib)₂-OMe. Even so, three unique conformations having origin transitions located at 37581 cm^{-1} , 37434 cm^{-1} , and 37629 cm^{-1} were identified based on hole-burning data, and are labeled A, B, and C, respectively. Based on the intensities of the S_0-S_1 origin transitions in the R2PI spectrum, conformer A appears to be the dominant conformation. The origin transition for conformer B appears only 2 cm^{-1} lower than the A origin in Z-(Aib)₂-OMe, suggesting a similar phenyl ring environment. Single-conformer UV spectra (lower traces of Fig. 6) were obtained for conformers A and B. (Note that the feature in the IR-UV HB spectrum of conformer B which appears at a frequency identical to the origin transition of conformer A is due to incomplete subtraction in the gated integrator.) However, a similar spectrum for conformer C was not available due to the weak intensity of



Fig. 6 R2PI (black) and IR-UV hole-burning spectra for conformers A (red) and B (blue) of Z-(Aib)_4-OMe. The hole-burning spectra were recorded with the IR hole-burning laser fixed at 3406 cm⁻¹ for A, and 3383 cm⁻¹ for B, marked with asterisks in Fig. 7(b).

its transitions. Instead, conformer C was identified by its possession of a unique IR spectrum.

RIDIR spectra of all three conformations collected in the C=O and NH stretching regions are shown in Fig. 7(a) and (b), respectively. As with the peptides examined above, calculated (M05-2X/6-31+G(d)), scaled (0.940), harmonic frequencies and infrared intensities for the assigned conformers are shown below each experimental spectrum. A good spectral match between a computed spectrum for a low-energy conformer and the experimental spectrum of conformer A of Z-(Aib)₄-OMe with four NH stretch transitions, occurring as two closely-spaced pairs, allows us to assign two lower-frequency H-bonded NH fundamentals (3383 cm⁻¹ and 3406 cm⁻¹) and two free NH stretch fundamentals (3464 cm⁻¹ and 3478 cm⁻¹).

The calculated spectrum for the assigned structure of conformer A associates a pair of C10-hydrogen-bonded NH groups with the low-frequency pair (3383 cm⁻¹ and 3406 cm⁻¹), classifying the structure as F-F-10-10. The H-bonded NH stretch fundamentals appear within a few cm⁻¹ of those found for the C10/C10 pair in the F-F-10-10 structure of Ac–Aib–Phe–Aib–NH₂ studied by Mons and co-workers.²⁴ This pattern of transitions indicates the first stages of a prototypical 3₁₀-helical structure in Z-(Aib)₄-OMe. Indeed, now all four Aib residues adopt local configurations characteristic of a full turn of a 3₁₀-helix (Table 1), with Ramachandran angles $(\phi, \psi) = (-61^{\circ}, -27^{\circ}), (-57^{\circ}, -22^{\circ}), (-57^{\circ}, -35^{\circ}), and (-48^{\circ}, -40^{\circ})$. The slight variation of the ϕ, ψ angles of the C-terminal Aib residue is not surprising in light of the fact that the C-terminal ester group, lacking the capacity to function as a hydrogen-bonding donor, does not support the formation of additional hydrogen bonds.

PCCP



Fig. 7 RIDIR spectra for conformers A, B, and C of Z-(Aib)₄-OMe in the (a) Amide I region and (b) NH stretch region. Asterisks indicate the transitions used for IR-UV hole-burning in Fig. 6. Calculated, scaled, harmonic, vibrational frequencies and infrared intensities for the assigned structures are shown as stick spectra below experiment.

Our calculated assignment for conformer A of Z-(Aib)₄-OMe is supported by the IR spectrum in the Amide I region shown in Fig. 7(a), which shows two closely spaced sets of transitions occurring at 1701 cm⁻¹/1707 cm⁻¹ and 1727 cm⁻¹/1739 cm⁻¹, as well as a single higher-frequency stretch at 1762 cm⁻¹. We note that these Amide I stretches (as well as those for conformers B and C of Z-(Aib)₄-OMe, discussed below) are shifted to higher frequency than similar oscillators observed for proteins in the solution phase.46,47 This is consistent with measurements made in the gas phase and is in accord with what has been reported previously for similar systems.^{22,48} Assignments in the Amide I region are complicated by the different functional groups (carbamate, amide, and ester) containing a C=O, and the coupling between C=O groups in the helix, which delocalizes the vibrational modes across more than one carbonyl. The labels in the figure identify the dominant contributor to each transition and the resulting hydrogen-bonded structure. These assignments match nicely with our established understanding of the Amide I region. Specifically, C=O groups involved in hydrogen bonding are shifted down in wavenumber from their otherwise free position. Note that ester and carbamate groups have C=O stretch fundamentals that are inherently higher in frequency by $\sim 50 \text{ cm}^{-1}$ than their amide C=O counterparts.²⁰ This difference in site frequencies of the various functional groups explains why a carbamate C=O appears at a higher

frequency than a free, interior C==O, even though it participates in a hydrogen bond.

The range of calculated frequencies and intensities for the Amide I normal modes shown in Fig. 7(a) reflects the variety of local electrostatic environments surrounding individual amide oscillators as well as the inter-amide coupling between local oscillators that gives rise to delocalized Amide I normal modes. Our gas phase spectra and their conformational assignments provide good tests for Hessian analysis, which seeks to extract local mode frequencies and inter-amide coupling constants.⁴⁸ Additionally, our solvent-free system could provide verification and/or calibration of the growing number of electrostatic maps, which have been developed for the purpose of Amide I band spectral prediction.⁴⁹

The RIDIR spectrum of conformer B in the NH stretch region is shown as the middle trace (blue) in Fig. 7(b). A good match with the calculated NH frequencies of an energy-minimized structure allows us to assign two higher-frequency, free NH stretch transitions, 3454 cm^{-1} and 3463 cm^{-1} , a single C10 H-bonded NH stretch transition, 3380 cm^{-1} , and a C7 hydrogen-bonded NH to the strongly shifted and broadened set of peaks near 3250 cm^{-1} . The calculated stick spectrum and the experimental spectrum match well, leading to an assignment of conformer B to the F-F-10-7 structure shown in Fig. 8(b). The higher-frequency free NH stretch, 3464 cm^{-1} , is due to the

N-terminal carbamate NH. The transition at 3250 cm^{-1} is assigned to an NH participating in a C7 H-bond, which is calculated to have an NH···O=C hydrogen-bond distance of just 1.94 Å, leading to its large shift of more than 120 cm⁻¹ below the C10 NH stretch. This structure is intriguing, as it shares much in common with the F-F-10-10 structure of conformer A, but formally exchanges a final C10 H-bond for a C7. We will return in the discussion to a further consideration of this structure and its relationship to the 3₁₀-helix.

The good match between the experimental IR spectrum of conformer B in the Amide I region, Fig. 7(a), with the calculated spectrum of the energy minimized structure confirms the assignment of an F-F-10-7 structure. The lowest frequency C==O stretch transition at 1681 cm⁻¹ belongs primarily to the strongest hydrogen-bonded C==O[2] which is involved in a tight C7 hydrogen-bonded ring. This oscillator is also coupled to the free C==O stretch of the adjacent C==O[3] at 1710 cm⁻¹. The band at 1729 cm⁻¹ is a free, interior C==O[1]. The C10 carbamate C==O[0] and free, ester C==O[4] are the two highest frequency Amide I fundamentals appearing at 1754 cm⁻¹ and 1770 cm⁻¹. As discussed previously, the carbonyl stretch of the carbamate is

a comparatively high frequency feature, despite its participation in hydrogen bonding.

Finally, the RIDIR spectrum for conformer C of Z-(Aib)₄-OMe in the NH stretch region is shown in green in Fig. 7(b). The spectrum is unique, with a single free NH stretch at 3462 cm^{-1} , two NH groups engaged in intermediate-strength H-bonds, at 3355 cm⁻¹ and 3382 cm⁻¹, and an NH group in a strong H-bond at 3270 cm⁻¹. The structure tentatively assigned to conformer C is shown in Fig. 8(c), labeled as 14/11-7-7-F, with a predicted IR spectrum given in the stick diagram below the experimental spectrum. The transition occurring at 3270 cm^{-1} is due to the strong C7[NH1] H-bond. The pair of transitions at 3355 cm^{-1} and 3382 cm^{-1} are delocalized combinations of the C14/C11 bonded NH and the NH involved in the weaker of the two C7 rings. The Amide I region of conformer C, Fig. 7(a), is less well-resolved than Amide I regions of the other two conformers, having four broad bands centered around 1690 cm⁻¹, 1712 cm^{-1} , 1728 cm^{-1} , and 1752 cm^{-1} . The missing fifth band is likely hidden under the breadth of one of these other, much broader bands. The observed transitions are consistent with the stick diagram predicted for the assigned structure, which predicts



Fig. 8 Assigned structures for conformers A, B, and C of Z-(Aib)₄-OMe. Both the Schellman and non-Schellman versions of conformer A have been shown. The dotted circle indicates the location of the Schellman helix reversal. Hydrogen-bonding interactions are shown with dotted red lines. Structural families, relative energies (M05-2X/6-31+G(d), kJ mol⁻¹), and dipole moments (D) are listed. Non-amide hydrogen atoms have been removed for clarity.

the major C=O groups responsible for each absorption shown in the figure.

IV. Analysis and discussion

The goal of this study is to evaluate the structural preferences of a series of Aib-based homopeptides of increasing size in the gas phase to determine whether their 3_{10} -helix forming propensity carries over from solution to the gas phase. We seek to extract and evaluate the spectroscopic signatures of the 3_{10} -helix, if it is formed, identify possible competing structures, and understand how the Aib residue influences them.

Our Aib homopeptides (Z-Aib-OH, Z-(Aib)₂-OMe, and Z-(Aib)₄-OMe) vary in their propensity to adopt dihedral angles in the 3_{10} -helical region of the Ramachandran plot ($\phi \sim -57^{\circ}$, $\psi \sim -30^{\circ}$), and in their capacities to form a C10 hydrogen bond (*i.e.*, $i \rightarrow i + 3$ hydrogen bond of a Type III β -turn) that is the structural hallmark of the 310-helix. For example, while the Z-Aib-OH molecule is incapable of forming a C10 intramolecular hydrogen bond, a minor conformer of Z-Aib-OH is seen to adopt Ramachandran angles that support such hydrogen bonds in larger members of the series. Conformer A of Z-(Aib)₂-OMe also displays a set of dihedral angles for the first Aib residue that are near those of a 3_{10} -helix, $(\phi, \psi = -57^{\circ}, -30^{\circ})$, even though it also lacks the backbone length necessary to make a fully formed turn of the 310-helix. Indeed, the (ϕ, ψ) angles of the C-terminal Aib of conformer A of Z-(Aib)₂-OMe are fully extended ($\phi \sim -180^\circ, \psi \sim -180^\circ$). Finally, the Z-(Aib)₄-OMe molecule, large enough to allow two intramolecular C10 hydrogen bonds, adopts a clear 310-helix as its dominant conformer, in which two C10 hydrogen bonds are formed, and all four residues adopt dihedral angles characteristic of the 310-helix (Table 1). Competing structures that incorporate C7, C11, and even C14 hydrogen-bonded rings form architectures classified as stalled or mixed helices, which make minor contributions to the R2PI spectrum. Even so, it is clear that in the gas-phase, free from intermolecular stabilization and solvent effects, Aib is a robust 310-helix former.

A. Energetics analysis

An energy level diagram comparing the calculated structures for Z-Aib-OH is shown in the left-most column of Fig. 9. The structures assigned for conformers A and B are labeled. The limited backbone length of Z-Aib-OH prohibits the formation of more complex hydrogen-bonding interactions, and all but four of the calculated structures fall within a 13 kJ mol⁻¹ energy window, containing only free OH groups and free or C5 hydrogenbonded NH groups. The four structures falling outside this window are all at least 35 kJ mol⁻¹ above the calculated minimum energy structure and possess hydrogen bonding interactions between either the NH group or the OH and the π -cloud on the phenyl ring. These geometries induce substantial strain on the backbone. Conformer A, with a single C5 intramolecular hydrogen bond, was found to be the global minimum. Conformer B, with no intramolecular hydrogen bonding, is only



Fig. 9 Energy level diagram for the relative, zero-point corrected energies $(kJ \text{ mol}^{-1})$ of the conformational minima of Z-Aib-OH (left) and Z-(Aib)₂-OMe (right) calculated at the M05-2X/6-31+G(d) level of theory. The conformational families of Z-Aib-OH are bracketed and labeled. The conformational families of Z-(Aib)₂-OMe are labeled below their respective columns.

4.5 kJ mol⁻¹ higher in energy. DFT calculations at the M05-2X/6-31+G(d) level of theory show two intervening conformers, plotted in the diagram, that are not observed experimentally. If the calculated energy ordering is correct, these two structures must have low energy barriers to isomerization to A, allowing their population to be removed during the cooling process in the expansion.

Not surprisingly, the energy level diagram for Z-(Aib)₂-OMe, displayed on the right-hand side of Fig. 9, shows a larger number of low-energy structures than the diagram for Z-Aib-OH. Structural families are grouped by H-bonded features into columns; individual structures are ordered according to relative energy. A significant number of the calculated structures contain one free NH group and one hydrogen-bonded NH group. The observed H-bond types include: (i) a C5 ring formed between the NH and C=O groups on a given Aib residue, characteristic of an extended peptide backbone, (ii) a C7 ring in which an NH group makes a more traditional, end-on interaction with the C=O of the adjacent Aib residue, and (iii) in a few cases, a special C5 in which the NH group of the second Aib residue interacts with the lone-pair on the backside of the backbone, methoxy-group oxygen. There is a single structure, assigned to conformer C, which contains two C5 hydrogen-bonded NH groups in a fully-extended peptide backbone geometry. In this sense, conformer C is the sole member of another conformational family (5-5), and could rightfully have its own column in the diagram, but has been grouped with the F-5 structures for efficient space utilization.

Conformer A, an F-5 structure, is found to be the global minimum, with its first Aib residue having Ramachandran angles close to that of a 3_{10} helix (Table 1). It is noteworthy that

this structure, incorporating a single C5 ring, is lower in energy by 5.2 kJ mol⁻¹ than conformer C, the fully extended structure, with its two weak C5 H-bonds. This energy stabilization of A over C must be attributed, at least in part, to the preference of the first Aib residue for the turn-like geometry (Fig. 5(a)) observed in a 3_{10} -helix.

Conformer B, which is 7.2 kJ mol⁻¹ above the global minimum, represents the first example of a C5/C7 bifurcated double ring in which a single NH group interacts with the nearestneighbor C=O groups on either side of it. We have already noted that conformer C, the 5-5 structure, is the lowest (and only) example of a fully-extended structure. Thus, all three observed conformers of Z-(Aib)₂-OMe are the lowest-energy examples of three unique conformational families, suggesting that barriers between members of the same family are relatively low, but interfamily barriers are substantial enough that collisional cooling between them is incomplete. The large intensity of conformer A in the R2PI spectrum relative to B and C is consistent with B and C being higher in energy, and therefore having smaller populations. Admittedly, other factors beyond population can affect relative sizes of peaks in the R2PI spectrum (e.g., Franck-Condon activity, S₁ lifetime, photoionization cross section). Once again, the lack of spectroscopic evidence for conformers of the F-7 and F-F conformational manifolds computed to have energies lower than those of assigned conformers B and C suggests that structures may have small barriers for isomerization to conformer A.

Fig. 10(a) presents an energy level diagram for Z-(Aib)₄-OMe, the largest member of the present series of Aib homopeptides. Once again, the structures have been grouped into columns based on conformational/hydrogen-bonding family type. The first three columns have well-defined intramolecular hydrogen-bonding patterns; however the three columns to the right simply group the structures by the number of amide-amide H-bonds they possess.

The low-energy region of the energy level diagram for Z-(Aib)₄-OMe is remarkably sparse for a molecule this size, having only two competing structures within 10 kJ mol⁻¹ of the global minimum. Both the global minimum and the second lowest-energy (+3.5 kJ mol⁻¹) structures are F-F-10-10 structures characteristic of a 3_{10} -helix, with very similar calculated IR spectra that are nominally consistent with the observed spectrum of conformer A, whose S_0 - S_1 origin dominates the R2PI spectrum. Conformer B is assigned to the structure 3.9 kJ mol⁻¹ above the global minimum, which is the lowest energy member of the F-F-10-7 family. Above 10 kJ mol⁻¹, there are many structures with similar energies belonging to a wide variety of H-bonding architectures, only one of which was observed experimentally, and is assigned as conformer C.

Interestingly, the calculations predict a large number of structures, >11 kJ mol⁻¹ above the calculated global minimum, trapped in F-F-10-10 conformational minima (left-most column of Fig. 10(a)). A C10 backbone hydrogen bond between the carbonyl oxygen of residue "*i*" and the amide hydrogen of residue "*i* + 3" is the defining characteristic of a β -turn, for which there is a taxonomy including many variations of ϕ and ψ backbone angles for residues *i* + 1 and *i* + 2.⁵⁰ Conformer A of Z-(Aib)₄-OMe exhibits "Type III" β -turns characteristic of 3₁₀-helices, for which (ϕ_2 and $\phi_3 \sim -57^\circ$, ψ_2 and $\psi_3 \sim -30^\circ$)^{50,51} The higherenergy members of the F-F-10-10 manifold exhibit other classes of β -turn structures, differing in the dihedral angles along the incipient 3₁₀-helix, and in the strength of the C10 H-bonds that are formed.



Fig. 10 (a) Energy level diagram showing different structures of Z-(Aib)₄-OMe grouped by hydrogen-bonding family. Energies (kJ mol⁻¹) are relative to the calculated global minimum structure of Z-(Aib)₄-OMe. (b) Energy level diagram comparing the assigned structures for Z-(Aib)₄-OMe with their counterparts in Z-(Gly)₄-OMe and a selection of other low-energy Z-(Gly)₄-OMe structures. See text for further discussion. Energies (kJ mol⁻¹) on the left-hand column are relative to the calculated global minimum of Z-(Aib)₄-OMe or Z-(Gly)₄-OMe (as appropriate). Energies on the right-hand axis of (b) are relative to the F-F-10-10-(S) structure of Z-(Gly)₄-OMe.

Paper

The classically defined β -turn structure does not proscribe the ϕ , ψ angles of the *i* + 3 residue participating in a β -turn. The two lowest-energy F-F-10-10 conformers we calculated for Z-(Aib)₄-OMe exhibit ϕ , ψ angles for the C-terminal (*i.e.*, fourth) residue that belong to the 3₁₀-helical region of the Ramachandran map. The major difference between these two low energy structures is that the minimum energy structure inverts the ϕ , ψ angles of its C-terminal residue, while the structure 3.5 kJ mol⁻¹ higher in energy does not. That is, the minimum energy structure exhibits a change of handedness for its fourth residue, as compared to preceding three residues. This "helix reversal", known as a Schellman motif, is a common phenomenon at helix termini in protein structures and in the crystal structures of Aib oligomers.^{52–55}

The structure assigned to conformer A, with an energy $3.5 \text{ kJ} \text{ mol}^{-1}$ above the calculated global minimum, maintains its handedness along the whole length of the peptide backbone (*i.e.*, non-Schellman) and provides a slightly better spectroscopic fit to experiment. The calculated dipole moments of the Schellman and non-Schellman structures are nearly identical, differing by only 0.07 D. A comparison of these two calculated spectra with experiment is included in the ESI† (Fig. S1).

Elsewhere in the diagram, Schellman/non-Schellman (S/NS) structural pairs appear. For instance, the F-F-10-10 structure at +17.5 kJ mol⁻¹ differs from the global minimum by changes in the dihedrals about the first Aib residue (ϕ_1 , ψ_1). (The energetic cost for this rearrangement is quite high.) Its NS counterpart is also observed, and appears at +20.5 kJ mol⁻¹, as anticipated.

The conformational variations within the F-F-10-10 conformational manifold include deviations of ϕ and ψ of Aib[1] and/or Aib[4], and in the orientation of the phenyl ring of the Z-cap. These differences and combinations there-of have been annotated on the energy level diagram in Fig. 10(a).

The orientation of the phenyl ring can also have a significant impact on the DFT-computed energy of a given structure. In Fig. 10(a), structures annotated with "RF" incorporate a "ring flip" in which the phenyl ring is rotated by $\sim 180^{\circ}$ around the C-C(H2)-O-C(==O) dihedral angle of the Z cap, as compared to this dihedral's value for the rings in the two lowest energy structures. Such a geometry change comes at an energy cost of about 8 kJ mol⁻¹. The phenyl ring orientation may further destabilize the structure by adopting an anti-configuration in which the dihedral angle value falls between $\pm 120^{\circ}$ to $\pm 180^{\circ}$. This orientation raises the energy of a given structure by approximately 12 kJ mol⁻¹. The isolated nature of these ring adjustments has been tested by manually changing the ring orientation, reoptimizing the structure at the DFT level, and comparing the resultant energies. In these tests, it was observed that adjusting the ring in this way only impacted the zero-point corrected energy of a particular structure and did not impact the predicted spectrum to any significant extent.

The remaining annotations, not explicitly discussed here, all occur at the ends of the molecule and primarily involve atoms that do not participate in intramolecular hydrogen bonding. By considering these geometry rearrangements individually and in concert with each other, the large number of unique structures within the F-F-10-10 family can be understood. Conformer B is assigned as an F-F-10-7 structure. The F-F-10-7 structural manifold emerges from the F-F-10-10 structures by distortions of the dihedral angles in the third Aib residue away from their archetypal 3_{10} -helical values. There are nine structures within the first 25 kJ mol⁻¹ that belong to this F-F-10-7 family. The lowest energy F-F-10-7 structure (*i.e.*, conformer B) is almost 10 kJ mol⁻¹ more stable than the next structure in this family. Just as in the F-F-10-10 structures, the other members of the F-F-10-7 family can be accounted for with combinations of ring flips, reversals of the Aib[3] Ramachandran angles, and modifications of ψ_4 .

Conformer C is tentatively assigned to a compactly folded, mixed-helix structure containing three intramolecularly-hydrogenbonded NH stretch transitions and a free NH stretch. This conformer appears near the top of Fig. 10(a), in a column depicting the energy levels of conformations exhibiting 3 or 4 intramolecular hydrogen bonds. (The final two columns of Fig. 10(a) indicate the energies of conformations found to contain one or two intramolecular hydrogen bonds.) The structure assigned to conformer C, a 14/11-7-7-F conformation, is shown in Fig. 8(c). While the fit between experiment and theory in Fig. 7 is reasonable, the assigned structure is 24 kJ mol⁻¹ above the global minimum, and keeps this assignment tentative. The combination of laser desorption with rapid collisional cooling in the supersonic expansion may be trapping a small portion of the population in this conformational well. Assignment of such a high-energy structure, while certainly uncommon, is not unprecedented. In fact, Dean et al. made a similar argument to justify the assignment of a high energy conformation in their work on Z-(Gly)3-OH.22

B. Comparison of Z-(Aib)_n with Z-(Gly)_n

A natural point of comparison exists between the present set of Aib homopeptides and their corresponding glycine (Gly) counterparts. One point of commonality between the two is that, unlike chiral amino acids, where formation of secondary structures such as helices occurs with a particular handedness (e.g., L-amino acids preferentially form right-handed helices).50,56,57 Aib and Gly are both achiral, and therefore have allowed Ramachandran angles in which $(+\phi, +\psi)$ and $(-\phi, -\psi)$ are equivalent. Perhaps more importantly, Gly, with its two hydrogens, is the most flexible of the amino acids, with allowed Ramachandran angles that span the largest range of any of the amino acids.⁵⁸ By contrast, Aib, with its two methyl groups, shows much more restricted (ϕ, ψ) preferences localized around $(+57^{\circ}, +30^{\circ})$ and $(-57^{\circ}, -30^{\circ})$, and readily forms 310-helices in solution. Comparing the inherent conformational preferences of Aib and Gly homopeptides in the gas phase gives unique insight regarding the effects of dimethyl substitution, particularly in light of the accumulating macrodipole of the 310-helix that accompanies the spatial correlation of amide-group orientations.59

The structure observed for conformer A of Z-Aib-OH is strikingly similar to the assigned conformations for Z-Gly-OH in both its UV and IR spectroscopy.²² The limited backbone length of Z-X-OH, where X = Aib or Gly, provides very little conformational flexibility, and makes the extended C5 structure

PCCP

a highly favorable geometry. The similarity in the UV spectra between both molecules reflects their nearly identical phenyl ring environments. The frequencies of the OH and NH stretch oscillators also fall within a few cm⁻¹ of each other, again as a consequence of the similar structures. However, conformer B of Z-Aib-OH does adopt a distinct conformation. The increased steric impact of dual methyl sidechains causes a distortion of the fully extended structure and results in a backbone geometry with dihedral angles characteristic of the 3₁₀-helix. With only a single Aib residue, the nascent ability of the dimethylated peptide to induce this geometry is observed in Z-Aib-OH. The analogous Gly-based structure does not experience this steric burden, and the second, minor conformer observed in Z-Gly-OH retains the same dihedral angles as the major conformer and simply experiences a reorientation of its aromatic ring.²²

The data for Z-(Aib)₄-OMe can be compared qualitatively to Z-(Gly)₃-OH, with the fourth amide group in Z-(Aib)₄-OMe having its NH group substituted for OH in Z-(Gly)₃-OH.²² Despite this nominal similarity, the dominant conformer of Z-(Aib)₄-OMe, the F-F-10-10, 3₁₀-helix, has no low-energy counterpart in Z-(Gly)₃-OH. In order for the Gly-based molecule to form a second C10 H-bond the cis-OH in the COOH group would have to reconfigure to trans, an energetically costly isomerization. Instead, the major conformer A of Z-(Gly)₃-OH has an 11-7-7-OH $\cdots \pi$ hydrogen-bonding pattern that is surprisingly similar to the 14/11-7-7-F pattern tentatively assigned to the minor conformer C in Z-(Aib)₄-OMe. The OH $\cdots \pi$ hydrogen bond in conformer A of Z-(Gly)3-OH would have its equivalent in an NH··· π H-bond in conformer C of Z-(Aib)₄-OMe, but formation of this π H-bond is blocked by the ester cap. Instead, the ester C=O[4] forms the second leg of a C14/C11 bifurcated double-ring with NH [1] in the Aib-based molecule.

One of the most striking results of the Z-(Gly)_n studies occurred in the largest members of the series, Z-(Gly)5-OH and Z-(Gly)5-NHMe. While polyglycine peptides are known to adopt PGI/PGII structures (extended structures having ϕ , ψ dihedral angles of -150° , 147° and -77° , 145° , respectively) in solution,⁶⁰ Dean *et al.* discovered that these molecules instead form 14/16 mixed helices in the gas phase. 22 Unlike $\alpha \text{-}$ or 3_{10}- helices, mixed helices incorporate H-bonds that include both N \rightarrow C terminus and C \rightarrow N terminus H-bonds that reduce the magnitude of the dipole moment which frequently accompanies helix formation.²² In the gas phase, the calculated energies of polyglycine PGII or 310-helices increase relative to the 14/16 mixed helix as the helix length grows. Dean et al. postulated that the accumulating macrodipole of the polyglycine helices (PGII (9 D), 310 (21 D)) destabilizes them in the non-polar environment of the gas phase relative to the low dipole moment of the mixed helix (μ = 2.8 D in Z-(Gly)₅-NHMe) leading to the preference for mixed helical structures.

The present work on Aib homopeptides shows the reverse to be true in Z-(Aib)₄-OMe. The F-F-10-10 structure, an emergent form of the 3₁₀-helix, is preferred over all alternatives, despite its sizable macrodipole of 13.7 D. We surmise on this basis that dimethyl substitution at C_{α} restricts the energetically allowed (ϕ, ψ) angles about each Aib residue to such an extent that 3₁₀-helix formation is observed in spite of its significant macrodipole. While a direct experimental comparison between Z-(Aib)₄-OMe and Z-(Gly)₄-OMe is not currently available, it is possible to calculate the relative energies of optimized structures for Z-(Gly)₄-OMe at the same level of theory, and thus gain insight to the changes in relative energy that accompany dimethyl substitution of glycine residues. In addition, in order to place these structures for Z-(Gly)₄-OMe into a more appropriate context, a conformational search was performed on Z-(Gly)₄-OMe, and a set of the low energy structures discovered by the force field level search was submitted to DFT geometry optimization and harmonic, vibrational frequency calculation. The results of these calculations are shown in the energy level diagram in Fig. 10(b).

Several important deductions follow. Perhaps most startling is that the F-F-10-10 structure that is the emergent form of a 310-helix and the global minimum in Z-(Aib)4-OMe is now more than 18 kJ mol⁻¹ above the lowest energy structure calculated from our cursory search of Z-(Gly)4-OMe. In contrast, the minimum energy Z-(Gly)₄-OMe structure is a 14-7-7-F structure, with its three hydrogen bonds pointing in both the N \rightarrow C terminus and $C \rightarrow N$ terminus directions, while those in the 3_{10} -helix point only from C \rightarrow N. It is worth emphasizing that we have not done an exhaustive search to find the global minimum structure in Z-(Gly)₄-OMe, so the true minimum may be even lower in energy than the 14-7-7-F structure. Nevertheless, the 14-7-7-F structure is an incipient form of the 14/16-helix observed in Z-(Gly)5-NHMe, which had a full H-bonding pattern of F-14-7-7-16-7 using the present labeling scheme. The 14-7-7-F structure already has elements of a mixed-helix in that the C14 ring constitutes an $i \rightarrow i + 3$ H-bond, while the C7 rings are $i + 2 \rightarrow i$. The dipole moment of this structure is 5.2 D, while the dipole moment of the F-F-10-10 polyglycine structure is 12 D, consistent with the previous deduction that polyglycines in the gas phase preferentially fold into structures that reduce the dipole moment.

Substitution of Gly for Aib also has dramatic effects on the relative energies of the observed conformers of Z-(Aib)₄-OMe. For instance, the energy ordering of the assigned F-F-10-10 and F-F-10-7 structures switches upon substitution of Gly for Aib. In the Aib-based structures, conformer B, which incorporates a C7 hydrogen bond instead of a second C10 (*i.e.*, F-F-10-7), is calculated to be approximately 4 kJ mol⁻¹ higher in energy than the global minimum (F-F-10-10) structure. When Gly residues compose the backbone instead, the F-F-10-7 structure becomes lower in energy than F-F-10-10(s) by ~6 kJ mol⁻¹, a relative energy shift of nearly 10 kJ mol⁻¹.

Even the high-energy 14-7-7-F structure tentatively assigned to conformer C (Fig. 10(b)) of Z-(Aib)₄-OMe undergoes a significant shift in relative energy upon replacement of all the Aib residues with glycine residues. This structure has an energy 24 kJ mol⁻¹ above the global minimum for the Z-(Aib)₄-OMe case. When Gly residues are present instead, the significantly decreased steric crowding from the side chains, lowers the energy of the Gly version of conformer C by approximately 17 kJ mol⁻¹ (relative to the F-F-10-10 structure), while maintaining the same overall structure. (There is a slight backbone dihedral adjustment

Table 2 Shown below are the structures, relative energies, H-bonding patterns, and relevant dihedral angles for Gly-based analogues of the assigned structures for $Z-(Aib)_4$ -OMe as well as the selected structures from our cursory conformational search on $Z-(Gly)_4$ -OMe. Structures having a direct Z-(Aib)_4-OMe counterpart are listed as S and NS for Schellman and non-Schellman, B for conf. B, and C for conf. C. Non-analogous Z-(Gly)_4-OMe structures are simply numbered based upon their energy ordering. All angles are given in degrees

Molecule	Relative energy (kJ mol ⁻¹)	H-bonding pattern	Ring orientation	ϕ (1)	ψ (1)	ϕ (2)	ψ (2)	φ (3)	ψ(3)	ϕ (4)	ψ (4)
Z-(Gly) ₄ -OMe (S)	+18	F-F-10-10 (S)	67	-66	-12	-69	-6	-109	25	69	14
Z-(Gly) ₄ -OMe (NS)	+22	F-F-10-10 (NS)	66	-65	-16	-70	-5	-96	1	-71	-13
Z-(Gly) ₄ -OMe (B)	+13	F-F-10-7	62	-69	-13	-73	-19	80	-74	-104	-176
Z-(Gly) ₄ -OMe (C)	+26	14-7-7-F	85	-90	14	80	-80	-64	150	53	41
Z-(Gly) ₄ -OMe #1	0	14-7-7-F	99	-81	66	83	-67	-93	2	137	171
Z-(Gly) ₄ -OMe #2	+5	F-11-7-7	-85	57	-143	-89	69	81	-54	-122	-168
Z-(Gly) ₄ -OMe #3	+11	14-7-7-F	-84	80	-78	-85	70	121	-23	-103	60
Z-(Gly) ₄ -OMe #4	+17	11-7-7-F	-101	81	-78	-83	54	148	-165	-86	114
Z-(Gly) ₄ -OMe #5	+22	11/14-7-7-π	-169	77	-75	-79	75	127	-161	-92	66
Z-(Gly) ₄ -OMe #6	+27	F-11-7-7 (OMe)	83	-55	140	91	-70	-81	54	135	13

that results in the breaking of the first C7 hydrogen bond. These dihedral angles are shown in Table 2.)

The observant reader will point out that the global minimum structure found by our cursory probe of the Z-(Gly)₄-OMe conformational space also adopts a 14-7-7-F geometry. Indeed, this is yet another instance in which adjustments to the dihedral angles along the backbone result in dramatic changes in energy. Upon substitution of Gly residues for Aib residues, the dihedral angles of conformer C change by no more than 20° at any given position. In fact most of the dihedrals change by less than 10° , indicating that the reduction in energy likely comes primarily from relief of the steric strain imposed by the methyl side chains. Comparing the dihedral angles of the higher energy 14-F-7-F (Gly-substituted Z-(Aib)₄-OMe conformer C) to the minimum energy 14-7-7-F structure of Z-(Gly)₄-OMe, on the other hand, reveals dihedral angles that differ to a much larger extent $(20^{\circ}-100+^{\circ})$ at each position. Therefore, we conclude that the reduction in energy is associated with adoption of a geometry that is inherently preferred by the backbone.

Finally, there is one aspect of the energy level diagram that is relatively conserved within the Aib \rightarrow Gly substitution; namely, the relative energy difference between the Schellman and non-Schellman structures for conformer A, which remains near 3 kJ mol⁻¹ between the two molecules. Thus, it appears the energy requirement for flipping the C-terminal residue from one handedness does not arise from the dispersive/London forces of Aib's methyl side chains, but instead may be due to the different electrostatic interaction of the polar ester group with the rest of the molecule. Indeed, the recent crystal structure of Z-(Aib)₃-Gly-OtBu, which is terminated with a glycine residue, also exhibits the C-terminal helix reversal of the Schellman motif.⁶¹

C. Comparison with crystal structures

Fig. 11 compares the calculated Schellman and non-Schellman structures for the F-F-10-10 conformer of Z-(Aib)₄-OMe with an available X-ray crystal structure of the same molecule taken from the work of Ranganathan *et al.*⁶² Along most of the peptide backbone, the overlay between the gas phase F-F-10-10 and the X-ray crystal structure is excellent, indicating that even in the presence of close neighbors in the solid state, intramolecular forces



Fig. 11 Schellman (a) and non-Schellman (b) versions of conformer A of Z-(Aib)₄-OMe overlaid with previously determined crystallographic structure for a Z-(Aib)₄-OMe oligomer possessing a 3_{10} -helical structure extracted from a unit cell. All hydrogen atoms have been removed to better show the backbone overlap. Hydrogen-bonding interactions have been indicated with dotted red lines.

dictate the structure of this molecule, and the crystal structure is essentially unperturbed from the gas phase minimum. The fact that the structure is stable regardless of its environment also speaks to the strong local structural preferences induced in the peptide backbone by the Aib residue, a rather remarkable result given the small size of the methyl groups.

One significant distortion is seen in the comparison between the non-Schellman structure and the crystal structure at the C-terminal end of the molecule. In the tentatively assigned (non-Schellman) gas-phase structure, the C-terminal Aib residue has dihedral angles that maintain the handedness of the helix, while in the crystal structure the handedness of the final residue is flipped. This difference is most clearly seen in the front view, shown in Fig. 11(b). In the crystal phase, the Schellman motif enables formation of a pair of intermolecular NH···C=O hydrogen bonds with each of two adjacent Aib oligomers within the unit cell. By aligning two 3_{10} -helical segments end to end, the free NH groups at the N-terminus of one oligomer interact with the free C=O groups located at the C-terminus of the adjacent molecule. The isolated molecules probed by our experiment removes the possibility for intermolecular hydrogen bonding between Aib oligomers. While the vibrational spectrum shows a slightly better fit with the standard 3_{10} -helix form of F-F-10-10, both structures are low in energy, with the Schellman version calculated to be the global minimum. Since the difference between the calculated infrared spectra for the two structures is subtle (see Fig. S1 in ESI†), if the energy difference is real, it is likely that the experimentally observed structure is the Schellman version, in which case the correspondence between gas phase and crystal is near-perfect.

V. Conclusion

The work presented here investigates the folding propensities of a series of capped Aib-based homopeptides, Z-Aib-OH, Z-(Aib)2-OMe, and Z-(Aib)₄-OMe. Single-conformation IR and UV spectra recorded under jet-cooled conditions in the gas phase are compared with harmonic, vibrational frequencies and infrared intensities calculated for the conformational minima of the molecules, leading to clear assignments. The achiral Aib residue is known to be a robust 310-helix former, with prototypical Ramachandran (ϕ, ψ) angles near $(-57^{\circ}, -30^{\circ})$ for a right-handed helix and $(+57^{\circ}, -30^{\circ})$ $+30^{\circ}$) for a left-handed helix, which are energetically equivalent under isolated conditions. These local, residue-level angle preferences are observed to some degree already in Z-Aib-OH and Z-(Aib)₂-OMe, even though these homopeptides are too short to form the C10 H-bonds that serve as a spectroscopic marker for 310-helix formation. This indicates that the dimethyl substitution is inherently directing enough to produce a preference towards these angles even when supporting hydrogen bonds cannot be formed. However, alternative structures are also observed in these shorter members of the sequence.

When the backbone is extended to four residues in Z-(Aib)₄-OMe, an F-F-10-10 structure dominates the spectrum, with its peptide backbone constituting more than one full turn of the 3_{10} -helix. This is marked by both the hydrogen-bonding pattern of the NH groups, F-F-10-10, and the ϕ and ψ angles near -57° and -30° , respectively for each of the Aib residues in the molecule. Calculations predict that the two lowest energy structures for Z-(Aib)₄-OMe are both F-F-10-10 in type: the fully right/lefthanded helix, and the helix which incorporates the Schellman motif at the last residue. A striking aspect of this result is that this strong energetic preference for the 3₁₀-helix formation occurs at the earliest stages it can, with only two C10 H-bonds in place. Furthermore, the 310-helix is seen even though the macrodipole of these structures is 13.7 D, the largest of any of the calculated structures, indicating that perhaps the lack of a polar solvent, which would stabilize the separation of charge associated with a large dipole moment in solution, is less consequential than was previously thought. This is the reverse of what is seen in Z-(Gly)_n-NHMe, where a mixed 14/16-helix was favored in part because of its small dipole moment, a seeming advantage in the non-polar environment of the gas phase. Since this macrodipole continues to grow with helix length, it will be interesting to see whether the gas-phase preference for 310-helix formation continues in longer members of this series, a task left for future work.

Acknowledgements

JRG, DMH, AOHC, KNB, and TSZ gratefully acknowledge support for this work from the National Science Foundation (NSF CHE-1465028). MCR gratefully acknowledges support for this work from the Lawrence Family Fellowship at Fairfield University.

References

- 1 D. A. C. Beck, D. O. V. Alonso, D. Inoyama and V. Daggett, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 12259–12264.
- 2 R. B. Best, D. de Sancho and J. Mittal, *Biophys. J.*, 2012, **102**, 1462–1467.
- 3 R. J. Moreau, C. R. Schubert, K. A. Nasr, M. Török, J. S. Miller, R. J. Kennedy and D. S. Kemp, *J. Am. Chem. Soc.*, 2009, **131**, 13107–13116.
- 4 K. O'Neil and W. DeGrado, Science, 1990, 250, 646-651.
- 5 S. Padmanabhan, S. Marqusee, T. Ridgeway, T. M. Laue and R. L. Baldwin, *Nature*, 1990, **344**, 268–270.
- 6 P. Lyu, M. Liff, L. Marky and N. Kallenbach, *Science*, 1990, 250, 669–673.
- 7 L. Pauling, R. B. Corey and H. R. Branson, *Proc. Natl. Acad. Sci. U. S. A.*, 1951, 37, 205–211.
- 8 C. A. F. Andersen and B. Rost, in *Structural Bioinformatics*, ed. J. Gu and H. Weissig, John Wiley & Sons, Inc., Hoboken, 2003, pp. 339–363.
- 9 J. Donohue, Proc. Natl. Acad. Sci. U. S. A., 1953, 39, 470-478.
- 10 C. Toniolo and E. Benedetti, *Trends Biochem. Sci.*, 1991, 16, 350–353.
- 11 C. Atilgan, Z. N. Gerek, S. B. Ozkan and A. R. Atilgan, *Biophys. J.*, 2010, **99**, 933–943.
- 12 S. H. Park and S. J. Opella, Protein Sci., 2007, 16, 2205-2215.
- 13 G. R. Marshall, E. E. Hodgkin, D. A. Langs, G. D. Smith, J. Zabrocki and M. T. Leplawy, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, 87, 487–491.
- C. Toniolo, G. M. Bonora, V. Barone, A. Bavoso, E. Benedetti,
 B. Di Blasio, P. Grimaldi, F. Lelj, V. Pavone and C. Pedone, *Macromolecules*, 1985, 18, 895–902.
- 15 A. Bavoso, E. Benedetti, B. Di Blasio, V. Pavone, C. Pedone, C. Toniolo and G. M. Bonora, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, 83, 1988–1992.
- 16 R. Banerjee and G. Basu, ChemBioChem, 2002, 3, 1263-1266.
- 17 Peptaibiotics: Fungal Peptides Containing alpha-Dialkyl alpha-Amino Acids, ed. C. Toniolo and H. Bruckner, VHCA, Wiley-VCH, Zurich, Weinheim, 2009.
- 18 H. Kim and M. Cho, Chem. Rev., 2013, 113, 5817-5847.
- 19 S. Krimm and J. Bandekar, in *Adv. Protein Chem.*, ed. J. T. E. C. B. Anfinsen and M. R. Frederic, Academic Press, 1986, vol. 38, pp. 181–364.
- 20 H. Maekawa, F. Formaggio, C. Toniolo and N.-H. Ge, J. Am. Chem. Soc., 2008, 130, 6556–6566.
- 21 T. Zeko, S. F. Hannigan, T. Jacisin, M. J. Guberman-Pfeffer, E. R. Falcone, M. J. Guildford, C. Szabo, K. E. Cole, J. Placido, E. Daly and M. A. Kubasik, *J. Phys. Chem. B*, 2014, **118**, 58–68.
- 22 J. C. Dean, E. G. Buchanan and T. S. Zwier, J. Am. Chem. Soc., 2012, 134, 17186–17201.

- 23 W. Chin, F. Piuzzi, J. P. Dognon, L. Dimicoli, B. Tardivel and M. Mons, J. Am. Chem. Soc., 2005, 127, 11900–11901.
- 24 V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel and M. Mons, J. Phys. Chem. A, 2007, 111, 7347–7354.
- 25 W. H. James III, E. E. Baquero, S. H. Choi, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2010, **114**, 1581–1591.
- 26 G. Meijer, M. S. d. Vries, H. E. Hunziker and H. R. Wendt, *Appl. Phys. B: Photophys. Laser Chem.*, 1990, **51**, 395–403.
- 27 J. C. Dean, PhD thesis, Purdue University, 2014.
- 28 D. M. Lubman, C. T. Rettner and R. N. Zare, J. Phys. Chem., 1982, 86, 1129–1135.
- 29 D. H. Levy, Annu. Rev. Phys. Chem., 1980, 31, 197-225.
- 30 M. D. Morse, Exp. Methods Phys. Sci., 1996, 29, 21-47.
- 31 D. M. Lubman, Mass Spectrom. Rev., 1988, 7, 559-592.
- 32 T. S. Zwier, J. Phys. Chem. A, 2001, 105, 8827-8839.
- 33 R. H. Page, Y. R. Shen and Y. T. Lee, *J. Chem. Phys.*, 1988, **88**, 4621–4636.
- 34 P. K. Weiner and P. A. Kollman, *J. Comput. Chem.*, 1981, 2, 287–303.
- 35 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440–467.
- 36 M. J. Frisch, G. W. Trucks, 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, Jr., 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, N. J. 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, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian 09, Gaussian, Inc., Wallingford CT, 2009.
- 37 Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, 3, 289–300.
- 38 P. S. Walsh, R. Kusaka, E. G. Buchanan, W. H. James III, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2013, **117**, 12350–12362.
- 39 J. R. Gord, P. S. Walsh, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. B*, 2014, 8246–8256.

- 40 R. Kusaka, D. Zhang, P. S. Walsh, J. R. Gord, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2013, **117**, 10847–10862.
- 41 A. M. Dwivedi, S. Krimm and B. R. Malcolm, *Biopolymers*, 1984, 23, 2025–2065.
- 42 J. Li and Y. Sha, *Molecules*, 2008, **13**, 1111–1119.
- 43 V. Santagada, F. Fiorino, E. Perissutti, B. Severino, V. De Filippis, B. Vivenzio and G. Caliendo, *Tetrahedron Lett.*, 2001, 42, 5171–5173.
- 44 E. Gloaguen, R. Pollet, F. Piuzzi, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2009, **11**, 11385–11388.
- 45 W. Chin, F. Piuzzi, I. Dimicoli and M. Mons, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1033–1048.
- 46 E. Vass, M. Hollosi, F. Besson and R. Buchet, *Chem. Rev.*, 2003, **103**, 1917–1954.
- 47 D. F. Kennedy, M. Crisma, C. Toniolo and D. Chapman, *Biochemistry*, 1991, **30**, 6541–6548.
- 48 E. G. Buchanan, W. H. James III, S. H. Choi, L. Guo, S. H. Gellman, C. W. Müller and T. S. Zwier, *J. Chem. Phys.*, 2012, **137**, 094301.
- 49 N. Sengupta, H. Maekawa, W. Zhuang, C. Toniolo, S. Mukamel, D. J. Tobias and N.-H. Ge, *J. Phys. Chem. B*, 2009, **113**, 12037–12049.
- 50 M. De Poli, M. De Zotti, J. Raftery, J. A. Aguilar, G. A. Morris and J. Clayden, *J. Org. Chem.*, 2013, **78**, 2248–2255.
- 51 C. M. Venkatachalam, *Biopolymers*, 1968, 6, 1425-1436.
- 52 C. Schellman, Univsersity of Regensburg: Regensburg, West Germany, 1979.
- 53 R. Aurora and G. D. Rose, Protein Sci., 1998, 7, 21-38.
- 54 S. Datta, N. Shamala, A. Banerjee, A. Pramanik, S. Bhattacharjya and P. Balaram, J. Am. Chem. Soc., 1997, 119, 9246–9251.
- 55 S. J. Pike, J. Raftery, S. J. Webb and J. Clayden, Org. Biomol. Chem., 2014, 12, 4124–4131.
- 56 T. E. Creighton, *Proteins: Structure and Molecular Properties*, 2nd edn, 1993.
- 57 R. A. Brown, T. Marcelli, M. De Poli, J. Sola and J. Clayden, Angew. Chem., Int. Ed., 2012, 51, 1395–1399.
- 58 R. J. Read, P. D. Adams, W. B. Arendall, III, A. T. Brunger, P. Emsley, R. P. Joosten, G. J. Kleywegt, E. B. Krissinel, T. Luetteke, Z. Otwinowski, A. Perrakis, J. S. Richardson, W. H. Sheffler, J. L. Smith, I. J. Tickle, G. Vriend and P. H. Zwart, *Structure*, 2011, **19**, 1395–1412.
- 59 W. G. Hol, Prog. Biophys. Mol. Biol., 1985, 45, 149-195.
- 60 S. Bykov and S. Asher, J. Phys. Chem. B, 2010, 114, 6636-6641.
- 61 R. Gessmann, H. Bruckner and K. Petratos, *Acta Crystallogr.*, 2014, C70, 1046–1049.
- 62 D. Ranganathan, S. Kurur and I. L. Karle, *Biopolymers*, 2000, 54, 249–261.