Supporting Information

Sequence dependent folding motifs of the secondary structures of Gly-Pro and Pro-Gly containing oligopeptides

Satish Kumar,[†] Kshetrimayum Borish,[†] Sanjit Dey, [†] Jayashree Nagesh[*#], and Aloke Das[*†]

[†]*Department of Chemistry, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pashan, Pune-411008*

[#]*Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore-560012, India*

Table of contents **Pages**

1. Synthesis procedures and characterization of the peptides S3-S8

1.1 Synthesis method S3-S4

1.2 ^{1}H NMR Characterization S5-S6

1.3 HRMS spectra of peptides S7-S8

2. Solution phase Spectroscopy methods S9-S11

2.1. Solution phase IR spectroscopy S9

2.2. Solution phase NMR spectroscopy S9

2.3. 2D-NMR spectroscopy S9

2.4. NMR titration S9-S11

3. Gas phase spectroscopy methods S12-S13

4. Conformational search, calculated structures, relative energies, and vibrational frequencies of the lowest energy conformers of both the peptides and benchmarking of the DFT functionals S14-S40

4.1. Conformational search for Boc-Gly-DPro-NHBn-OMe and Boc-DPro-Gly-NHBn-OMe peptides S14-S16

4.2. Performance of B3LYP-D3/def2-TZVPP, B97-D3/def2-TZVPP, ωB97X-D/def2-TZVPP, and M06-2X/6-311++G(2d,2p) in predicting energetics and IR frequencies of the observed conformers of the peptides in the experiment S17-S38

4.3. Energy-ordered conformers obtained using solvent calculations S39-S40

5. CSD study of both the sequences S41

6. X-ray single Crystal structure S42-S43

7. References S44

1. Synthesis procedures and characterization of the peptides

1.1 Synthesis method

Both the peptides Boc-^{D}Pro-Gly-NHBn-OMe and Boc-Gly-^{D}Pro-NHBn-OMe were synthesized by standard synthetic procedure reported in the literature.[1]

Synthetic procedure of Boc-D-Pro-Gly-NHBn-OMe: Commercially available Boc-D-Pro-COOH was coupled with N-Hydroxysuccinimide (NHS) in the presence of N,N'-Dicyclohexylcarbodiimide (DCC), tetrahydrofuran (THF) in a round bottom flask kept at ice condition and left at room temperature for 12 h. The product was then coupled with Glycine in the presence of $NaHCO_3$, THF and then acidified with 10% HCl and this resulted the formation of Boc-^{D}Pro-Gly-OH. Boc-^{D}Pro-Gly-OH was then mixed with 4-Methoxybenzylamine in the presence of EDC.HCl, HOBt, DIPEA, and DMF at 0^0 C. The final product was received after 12 hours of reaction at room temperature. The final compound was purified by column chromatography using ethyl acetate and hexane as solvent in 65% yield. Synthetic procedure has been shown also in the Figure S1 provided below.

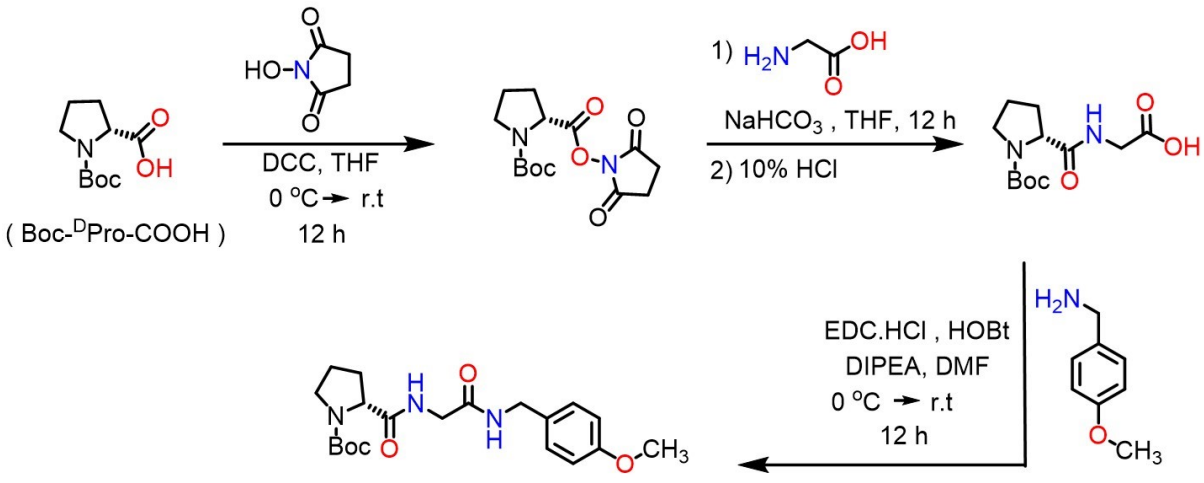


Figure S1. Synthetic Scheme of Boc-^{D}Pro-Gly-NHBn-OMe

Synthetic procedure of Boc-Gly-^{D}Pro-NHBn-OMe: Commercially available Boc-Gly-COOH was coupled with N-Hydroxysuccinimide (NHS) in the presence of N, N'-Dicyclohexylcarbodiimide (DCC), tetrahydrofuran (THF) in a round bottom flask kept at ice condition and then left at room temperature for 12 h. The product was then coupled with the Boc-^{D}Pro-OH in the presence of $NaHCO_3$, THF and then acidified with 10% HCl and this resulted the formation of Boc-Gly-^{D}Pro-OH. The product Boc-Gly-^{D}Pro-OH was then coupled with 4-Methoxybenzylamine in the presence of EDC.HCl, HOBt, DIPEA, DMF at 0^0 C. Afterwards, the reaction was maintained at room temperature and kept for 12 hours with continuation of the stirring. The final compound was purified by column chromatography using ethyl acetate and hexane as solvent in 72% yield. The synthetic procedure has also been depicted in Figure S2 provided below.

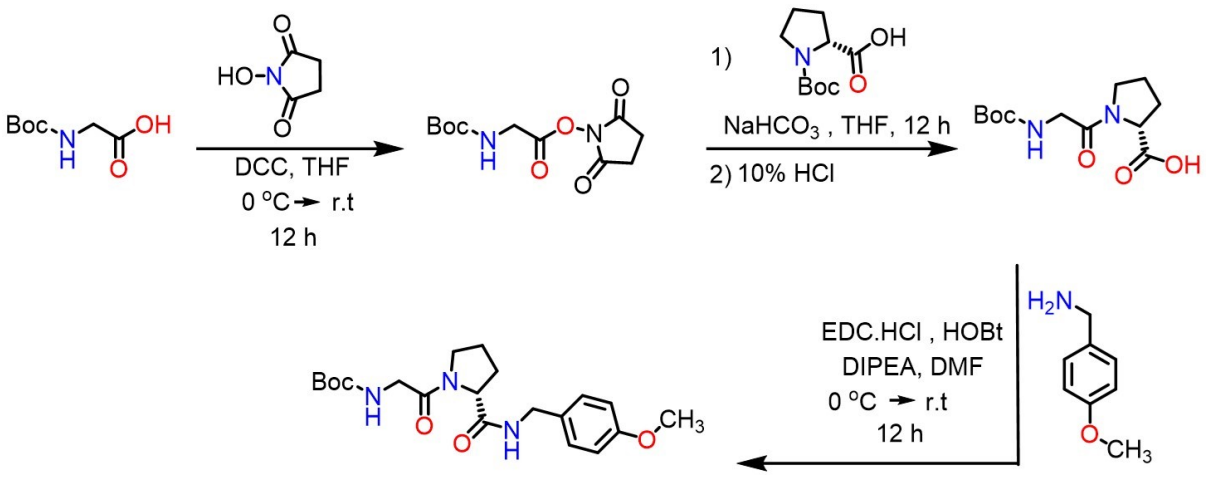
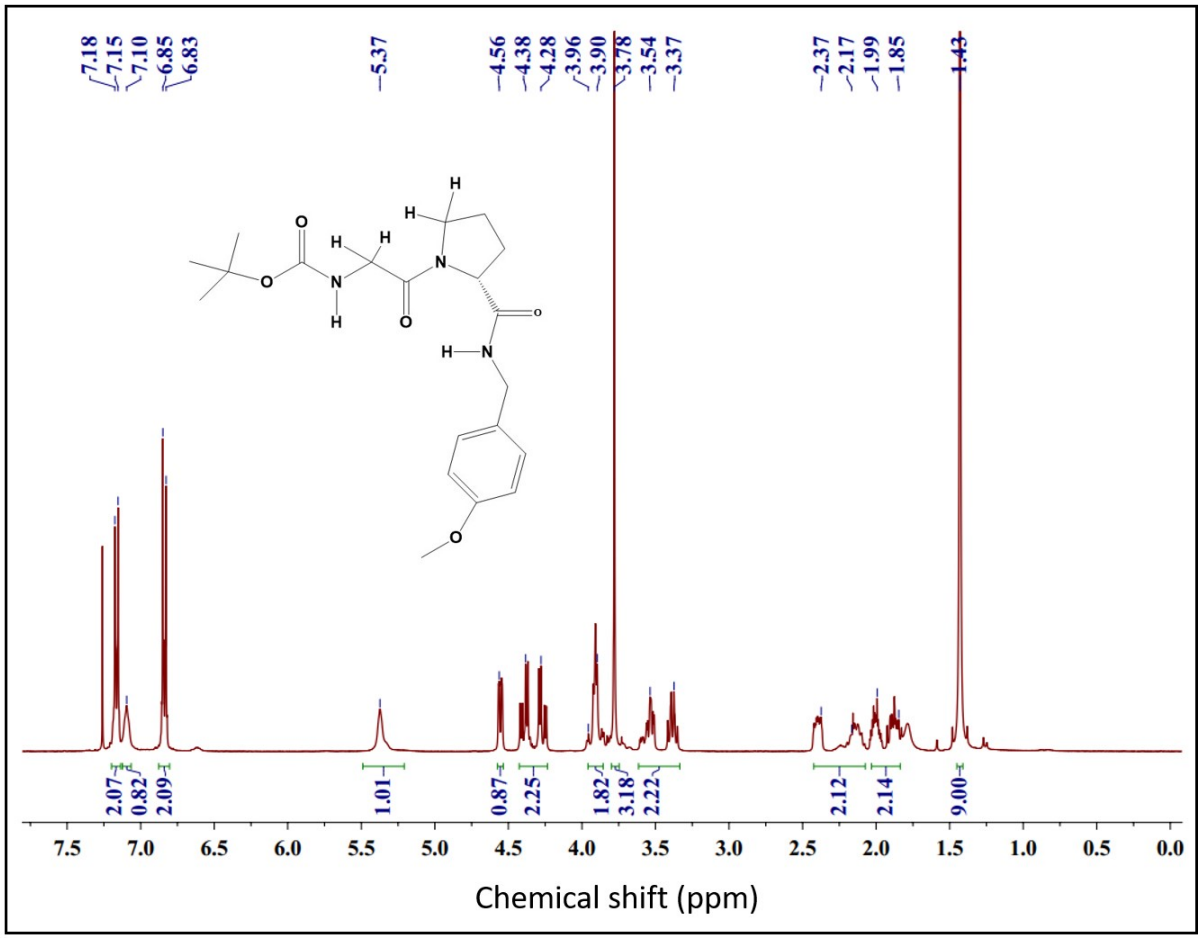


Figure S2. Synthetic Scheme of Boc-Gly-^{D}Pro-NHBn-OMe.

1.2 ^{1}H NMR Characterization

^{1}H NMR of Boc-^{D}Pro-Gly-NHBn-OMe (400 MHz, $CDCl_3$, 298.15 K): δ 6.80-7.50 (m,6H, NH_{Gly}, NH_{NHBn}, CH^{Ar}) 3.4-4.5 (m, 10H, CH_{Gly}, OCH_3, CH_{NHBn}, CH_{Pro}) 1.8-2.2 (m,4H, CH_{Pro}) 1.37 (s, 9H, CH_{Boc})

^{1}H NMR of Boc-Gly-^{D}Pro-NHBn-OMe (400 MHz, $CDCl_3$, 298.15 K): δ 6.80-7.50 (m,5H, NH_{NHBn}, CH^{Ar}), 5.37(s, 1H, NH_{Gly}) 3.4-4.5 (m, 10H, CH_{Gly}, OCH_3, CH_{NHBn}, CH_{Pro}) 1.8-2.2 (m,4H, CH_{Pro}) 1.43 (s, 9H, CH_{Boc})



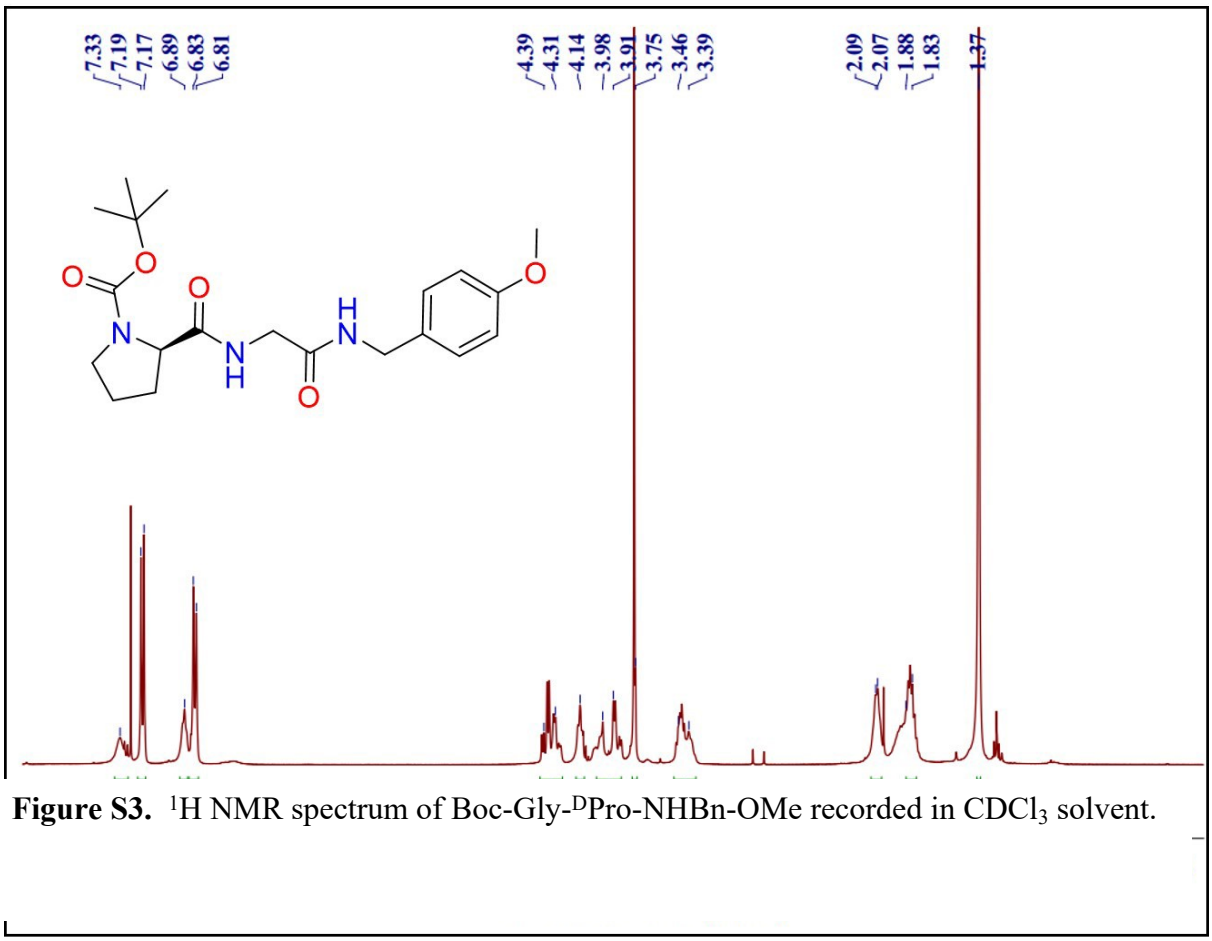
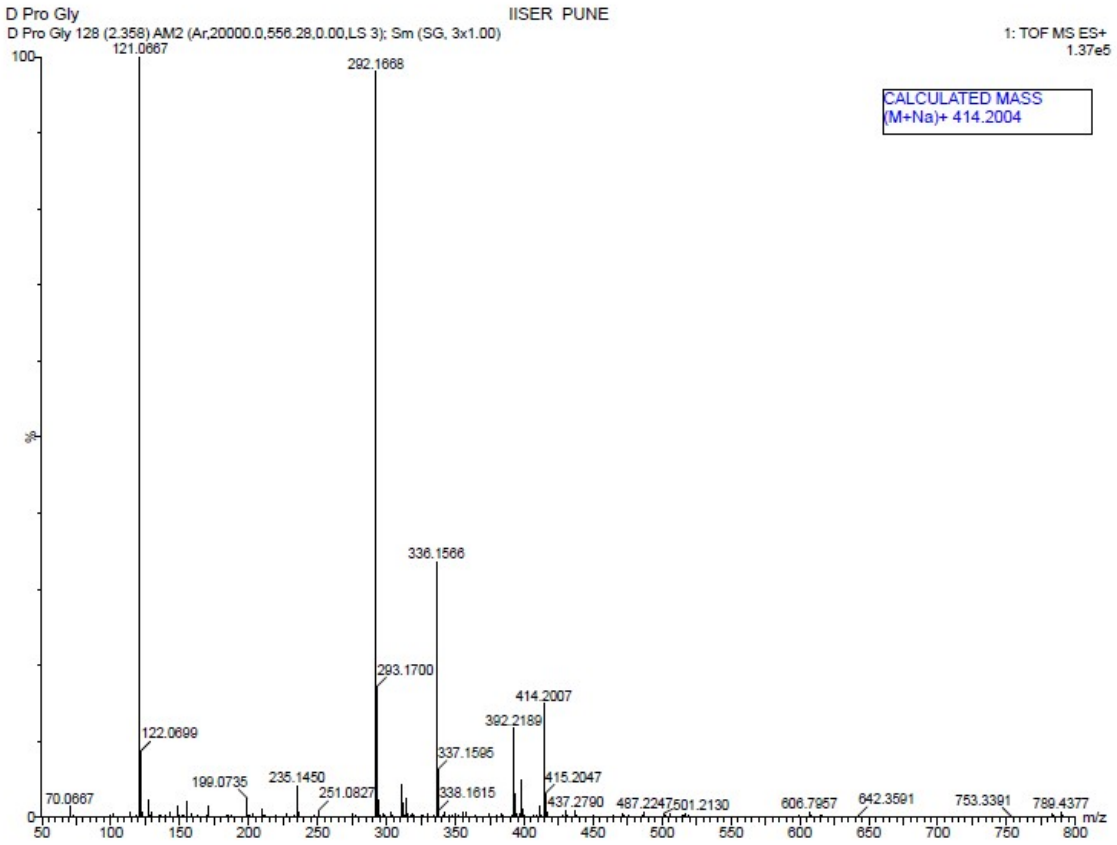


Figure S3. ^{1}H NMR spectrum of Boc-Gly-^{D}Pro-NHBn-OMe recorded in $CDCl_3$ solvent.

Figure S4. ^{1}H NMR spectrum of Boc-^{D}Pro-Gly-NHBn-OMe recorded in $CDCl_3$ solvent.



1.3 HRMS spectra of peptides

Figure S5. HRMS mass spectrum of Boc-^{D}Pro-Gly-NHBn-OMe.

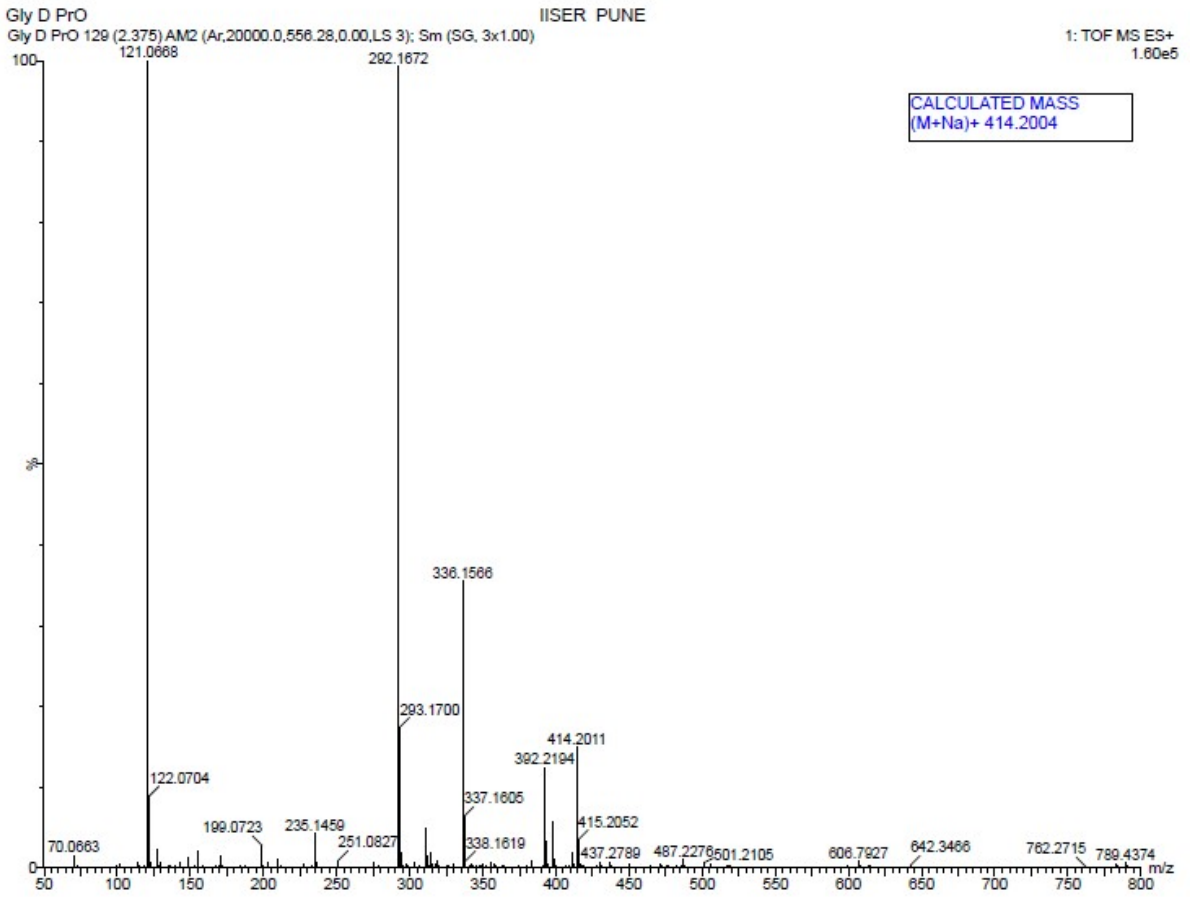


Figure S6. HRMS mass spectrum of Boc-Gly-^{D}Pro-NHBn-OMe.

2. Solution phase Spectroscopy methods

2.1. Solution phase IR spectroscopy.

Solution phase IR spectra of Boc-Gly-^{D}Pro-NHBn-OMe and Boc-^{D}Pro-Gly-NHBn-OMe peptides were measured at 293 K using Fourier-Transform IR spectrometer (Bruker Vertex 70). To record the IR spectra, the peptides are dissolved in $CDCl_3$ solvent and the concentration of the peptide solutions were maintained at 8 mM. The peptide solutions were put in a cell of 1 mm path length consisting of CaF_2 windows.

2.2. Solution phase NMR spectroscopy.

Both 1D and 2D ^{1}H NMR spectra of Boc-Gly-^{D}Pro-NHBn-OMe and Boc-^{D}Pro-Gly-NHBn-OMe peptides were measured in $CDCl_3$ solvent using a 400 MHz NMR spectrometer (Bruker-400). 1D ^{1}H NMR spectra were measured in $CDCl_3$ for characterization of the peptides synthesized in this work.

2.3. 2D-NMR spectroscopy.

Rotating frame Overhauser SpectroscopY (ROESY) of the two peptides were performed in $CDCl_3$ solvent using a 400 MHz NMR spectrometer (Bruker-400). The ROESY spectra reveal the correlations between the protons that are separated by few bonds but are spatially close. These spectra are useful to determine the structures of different conformers of the peptides.

2.4. NMR titration.

NMR titration with DMSO-d_6 by probing the chemical shift positions of the N-H groups in the peptides is performed to determine whether the N-H group is free or intramolecular hydrogen bonded. The chemical shift positions of the N-H groups were monitored by successive addition of 5 μL DMSO-d_6 in the $CDCl_3$ solution of the peptides. If there is significant downfield chemical shift or deshielding of the N-H proton by addition of DMSO-d_6, the N-H group in the

peptide is free and involved in intermolecular hydrogen bonding with DMSO. On the other hand, the change in the chemical shift of the N-H proton will be minimal or negligible on addition of DMSO-d_6 if the N-H group is already involved in intramolecular hydrogen bonding in the peptide.

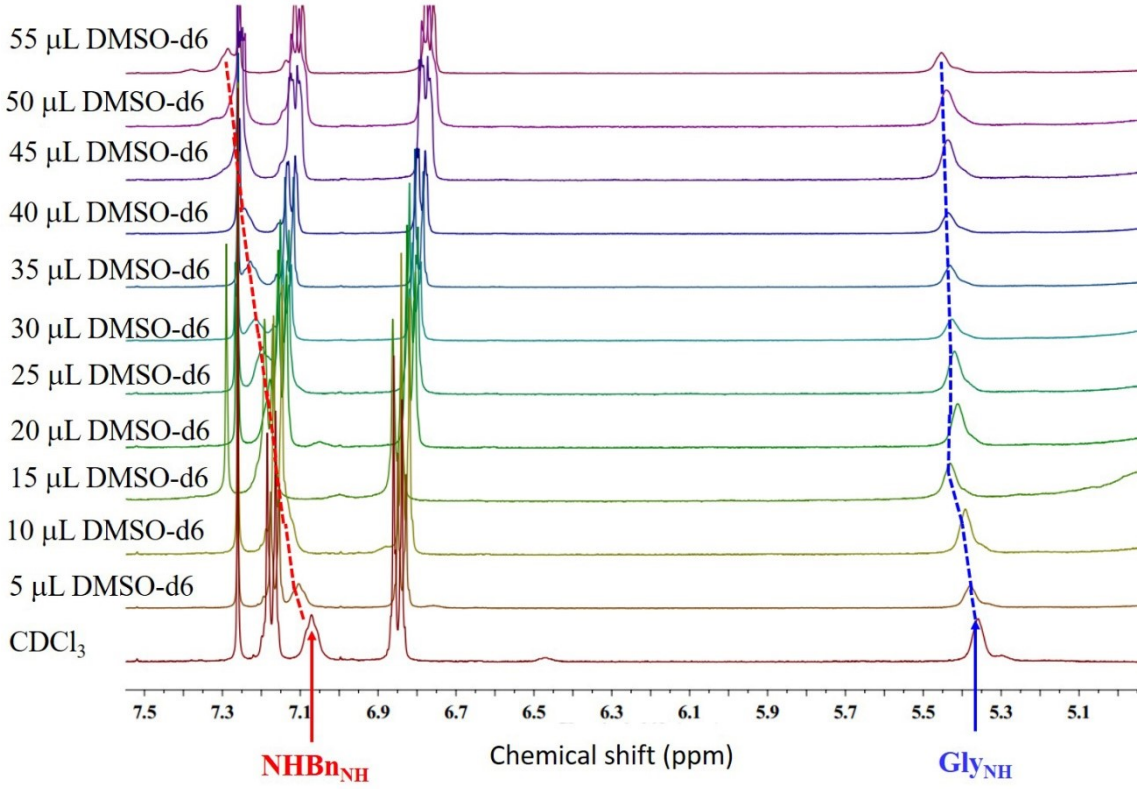


Figure S7. NMR titration of Boc-Gly-^{D}Pro-NHBn-OMe in $CDCl_3$ with stepwise addition of DMSO-d_6 by monitoring the chemical shift positions of the Gly and Bn N-H protons.

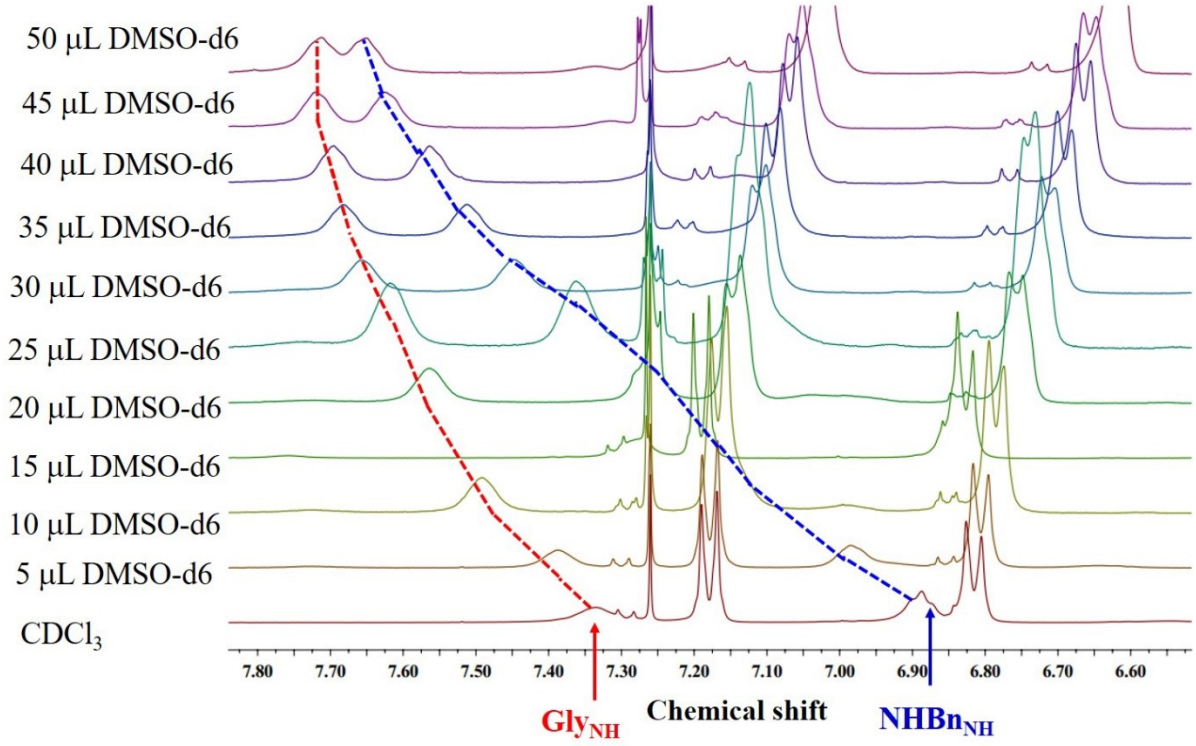


Figure S8. NMR titration of Boc-^{D}Pro-Gly-NHBn-OMe in $CDCl_3$ with stepwise addition of DMSO-d_6 by monitoring the chemical shift positions of the Gly and Bn N-H protons.

DMSO-d6 titration of Boc-Gly-^{D}Pro-NHBN-OMe sequence shows both NH are involved in intra-molecular hydrogen bond as chemical shift of both the NH upon addition of DMSO-d6 is very minimal.

In the case of Boc-^{D}Pro-Gly-NHBn-OMe glycine NH shows very less chemical shift compared to NHBn NH upon addition of DMSO-d6. It shows that in solution glycine NH is free whereas NHBn NH is involved in hydrogen bonding.

3. Gas phase spectroscopy methods.

The peptide molecules were brought into the vapor phase without any fragmentation using a laser desorption technique, which has been described in detail elsewhere.[2] A homogeneous mixture (2:1) of the peptide and graphite powder (Sigma Aldrich, size ~ 20 microns) was pressed in a hydraulic press with 2-3 tons of pressure to prepare a pellet of 6 mm diameter and 2 mm thickness. The pellet was put in a sample holder connected with a motorized assembly which was mounted with the vacuum chamber. The sample pellet was placed near the pulse valve in such a manner that it was away from the axis of the molecular beam by ~ 2 mm while the side edge of the pellet was maintained at a distance of ~ 1 mm from the face plate of the pulse valve. Second harmonic output (~ 0.6-0.7 mJ at 532 nm) of a Nd:YAG laser (Continuum, Minilite-I, 10 Hz, 10 ns) was mildly focused on the surface of the pellet which was near the pulse valve so that the desorption happened within 2-3 mm from the orifice of the pulse valve. The sample pellet was rotated at a speed of 0.5 mm/sec to avail the fresh spot of the pellet surface by different pulses of the laser and also to avoid the burning of the pellet surface.

The desorbed peptide molecules were enrouted in the supersonic expansion of Ar (~5 bar) carrier gas and internally cooled through enormous collision with carrier gas near the valve orifice. The cold peptide molecules were ionized by second harmonic output (0.2–0.3 mJ) of a tunable dye laser (ND6000, Continuum) pumped by frequency doubled output of an Nd:YAG laser (10 nanoseconds, 10 Hz, Surelite II-10, Continuum) and the ions were analyzed in the TOF mass spectrometer (Jordan TOF, USA). The details of the experimental setup were described elsewhere.[3, 4]

Mass-selected electronic spectra of the peptides were recorded by scanning the UV laser using one-color resonant 2-photon ionization (1C-R2PI) method. In this technique, the first photon of the laser excites the molecules to the first excited electronic state (S_1) while the second photon of the laser at the same wavelength ionizes the molecules. UV-UV hole-burning

spectroscopy was used to determine the presence of different conformers of the peptide in the experiment. In this method, one of the UV lasers (probe laser, 10 ns, 10 Hz, ND6000, Continuum) was fixed at the wavelength of one of the electronic bands in the R2PI spectrum and another pump UV laser (10 ns, 10 Hz, ND6000, Continuum) fired ~ 150 ns prior to the probe laser scanned in the region of the electronic spectrum. Both the lasers were in the counter-propagating direction and intersecting the molecular beam in the mutually perpendicular fashion. The depletion of the ion signal of the probe laser beam is observed due to the R2PI bands which belong to the same conformer of the peptide. The IR spectra of the peptides were measured using resonant ion-dip infrared spectroscopy (RIDIRS). In the RIDIRS, the UV laser wavelength was fixed at one of the electronic bands of a conformer and the IR laser preceding the UV laser by ~100 was tuned in the region of the N-H stretching region of the peptide. The IR spectra were obtained as a depletion in the ion signal when the IR laser frequency was matching with the vibrational frequency of the molecules. Finally, IR-UV hole-burning spectroscopy was used to determine the presence of different conformers of the peptide in the experiment. In this technique, the IR laser wavelength was fixed at one of the N-H frequencies of the peptides and the UV laser fired ~100 ns after the UV laser was scanned in the whole electronic spectral region of the peptide. The depletion in the UV ion signal is obtained for the R2PI bands arises from the same conformer. A tunable IR laser (Laser Vision, pulse energy ~ 4-5 mJ, resolution ~ 2.5 cm^{-1}) based on optical parametric oscillator (OPO)/ optical parametric amplifier (OPA) pumped by an unseeded Nd: YAG laser (Continuum, Surelite II-10, 10 nanoseconds, 10 Hz) was used for the IR-UV double resonance experiment.

4. Conformational search, calculated structures, relative energies, and vibrational frequencies of the lowest energy conformers of both the peptides and benchmarking of the DFT functionals

4.1. Conformational search for Boc-Gly-^{D}Pro-NHBn-OMe and Boc-^{D}Pro-Gly-NHBn-OMe peptides

4.1.1. Flowchart of process of conformational selection

The following flowchart describes the detailed conformational search using MMFF94 force field calculation as well as chemical intuition followed by quantum chemistry calculations of a large set of conformers at the HF/6-31G(d), M05-2X/6-31+G(d), B3LYP-D3/def2-TZVPP, B97-D3/def2-TZVPP, ωB97X-D/def2-TZVPP, and M06-2X/6-311++G(2d,2p) levels of theory. In fact, M05-2X, B3LYP-D3, B97-D3, and ωB97X-D level calculations are done for the Gly-Pro peptide for 68 structures up to energy ~ 25 kJ/mol while the same is done for the Pro-Gly peptide for 52 structures up to ~ 16 kJ/mol. Finally, M06-2X/6-311++G(2d,2p) level calculations are done for the Gly-Pro peptide for 16 structures up to energy ~ 14 kJ/mol and the same has been done for the Pro-Gly peptide for 16 structures up to energy ~ 8 kJ/mol.

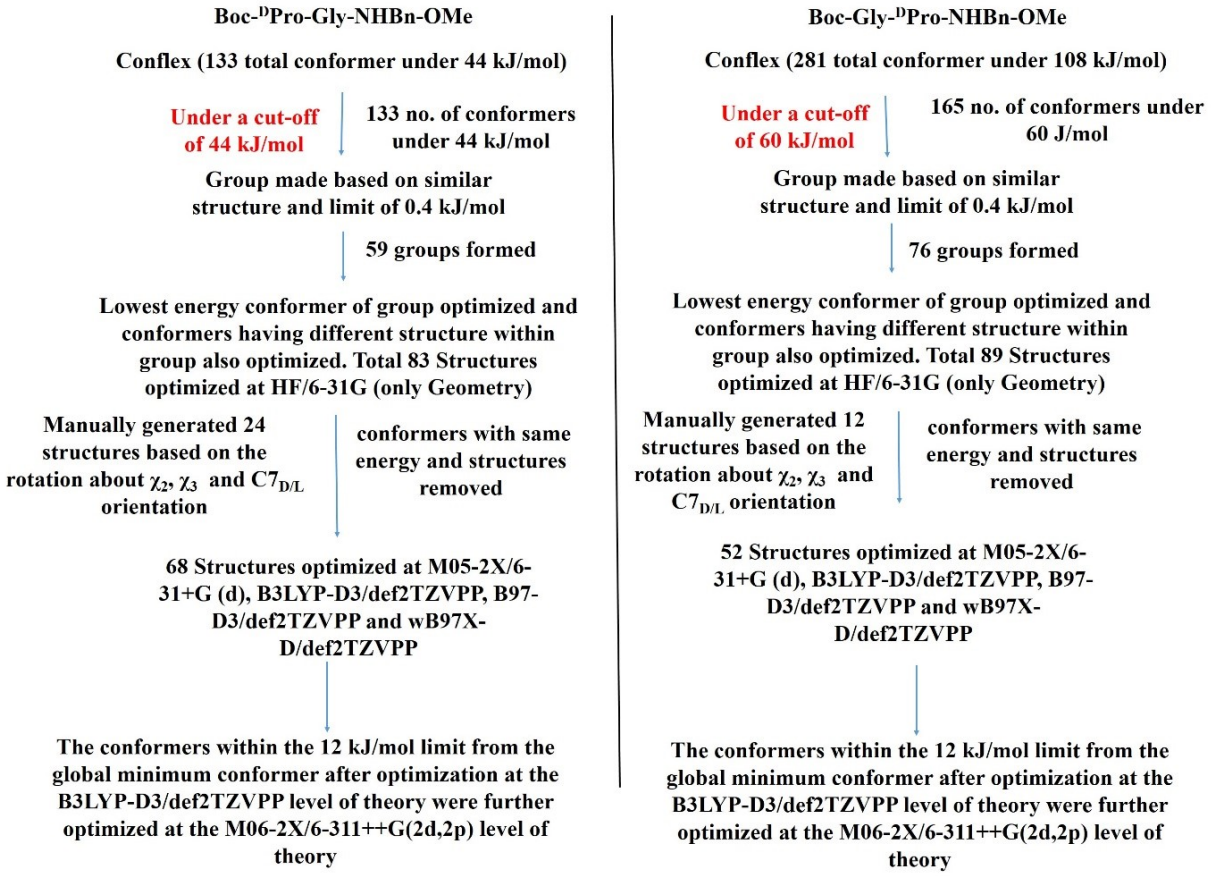


Figure S9. Conformational search procedure for Gly-Pro and Pro-Gly sequences using a combination of molecular mechanics and manual search methods.

4.1.2. Dihedral angles used for manual generation of the conformers of the Gly-Pro and Pro-Gly peptides

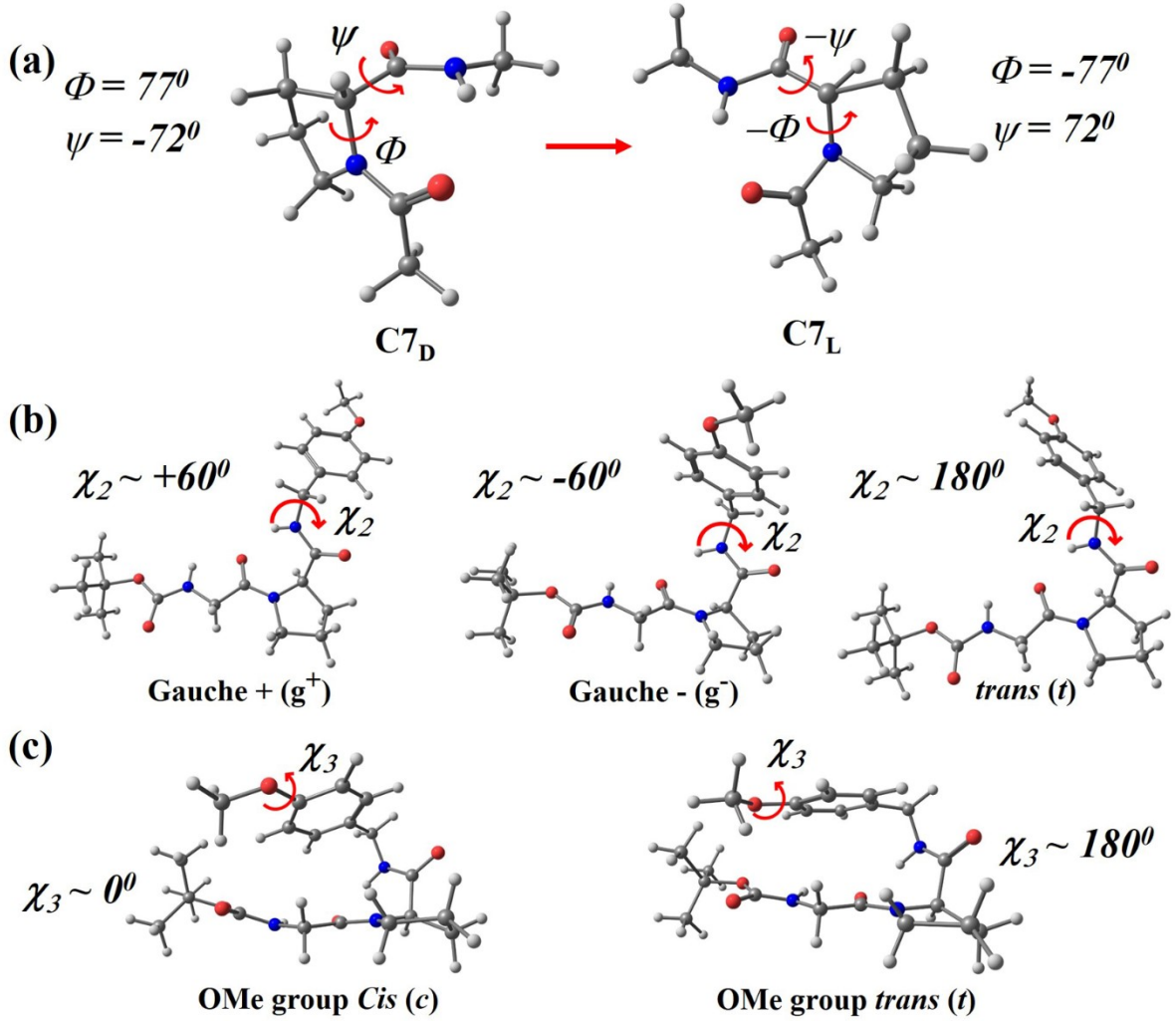


Figure S10. Conformers named based on (a) C7 D&L orientation (b) Bz group orientation and (c) OMe group orientation.

4.2. Performance of B3LYP-D3/def2-TZVPP, B97-D3/def2-TZVPP, ωB97X-D/def2-TZVPP, and M06-2X/6-311++G(2d,2p) in predicting energetics and IR frequencies of the observed conformers of the peptides in the experiment

The numbering of the conformers of the peptides (GP1….GPn or PG1….PGn) in terms of energy ranking is done at the M06-2X/6-311++G(2d,2p) level of theory as this level of calculation does better job in terms of the assignment of the observed conformers of both the peptides. The same numbered conformers are simply used for calculations of the energies and frequencies at other levels of DFT.

4.2.1. Energy landscapes of Gly-Pro peptide at different levels of theory

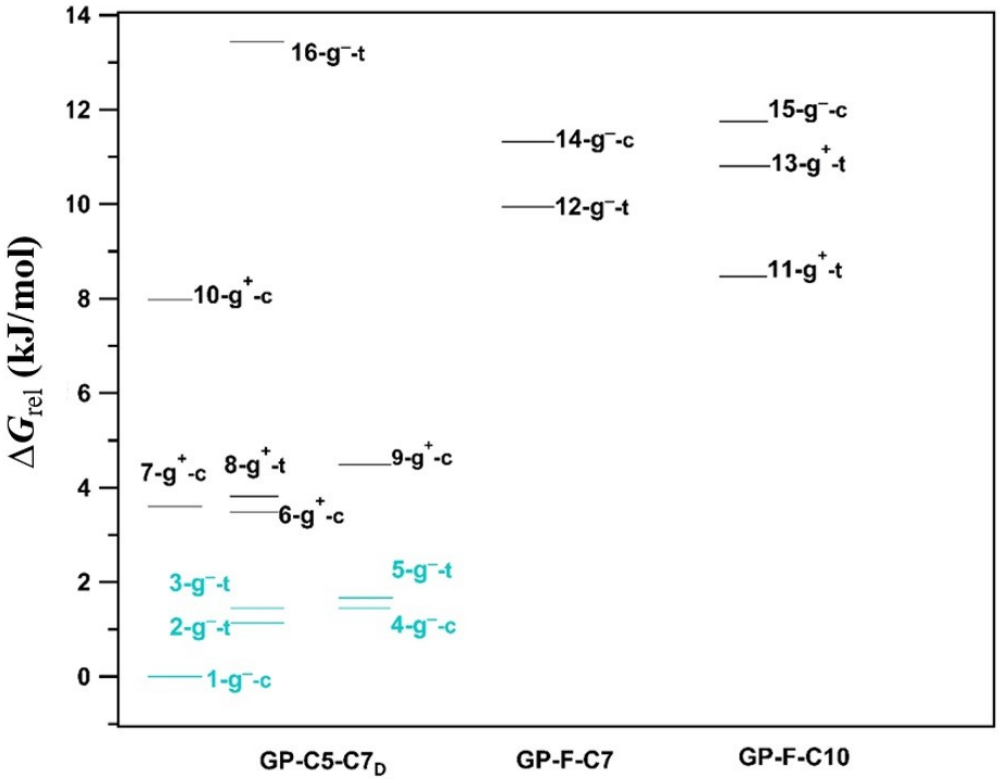


Figure S11. Energy landscape of a few low energy conformers of (a) Boc-Gly-^{D}Pro-NHBn-OMe calculated at 300 K at the M06-2X/6-311++G(2d,2p) level of theory. The color coded conformers are the predicted ones observed in the experiment. See the main text for the nomenclature.

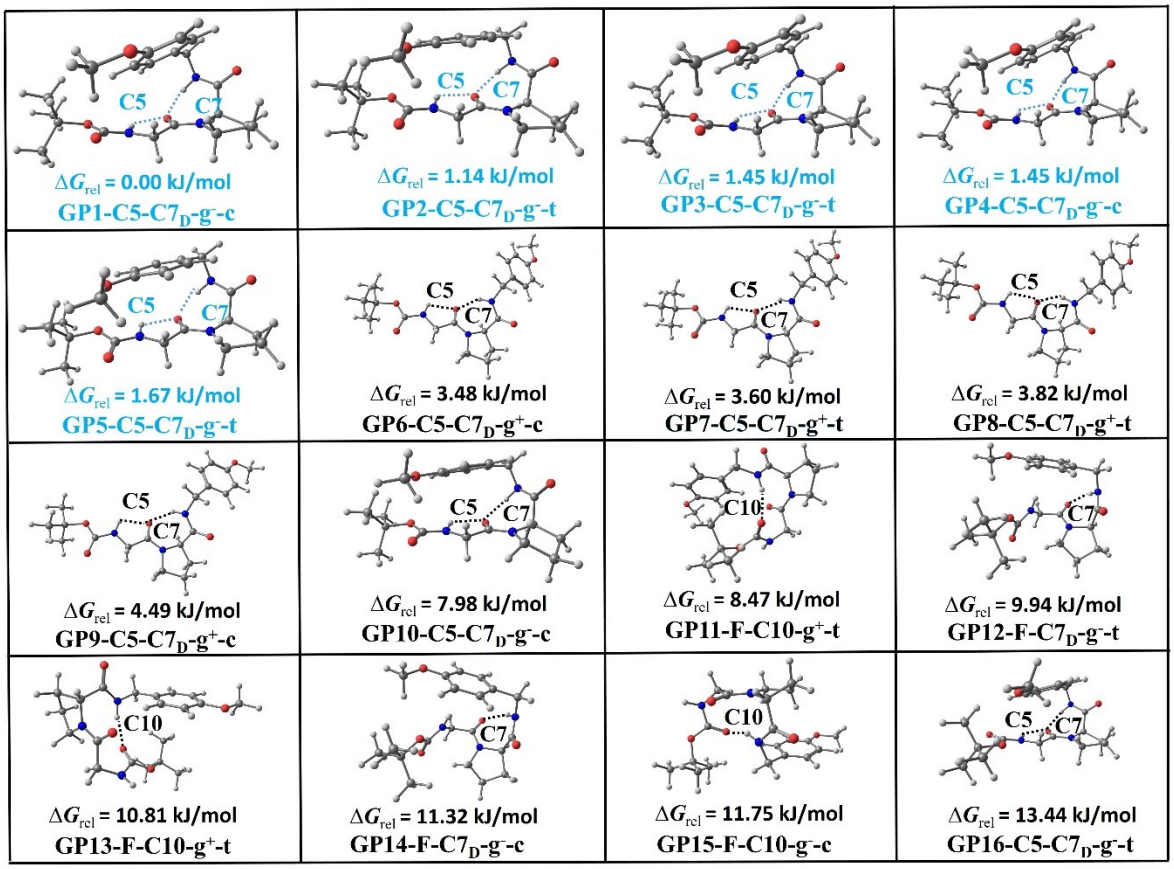


Figure S12. Structures of the 16 lowest energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe optimized at the M06-2X/6-311++G(2d,2p) level of theory. The color coded conformers are the predicted ones observed in the experiment.

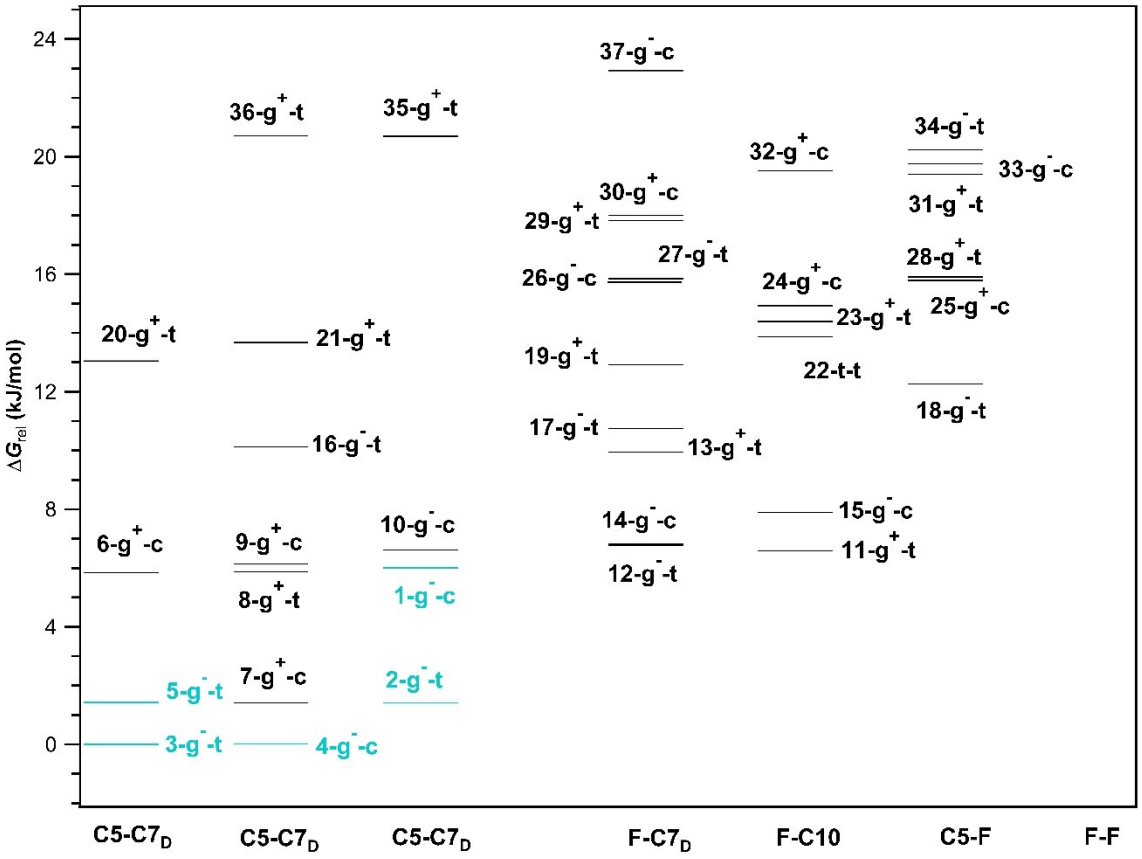


Figure S13. Energy landscape of Boc-Gly-^{D}Pro-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer calculated at 300 K at the B3LYP-D3/def2-TZVPP level of theory. The label C5-C7$_D$ is used thrice merely to show the close-lying energy levels in a distinct manner. See the main text for the nomenclature.

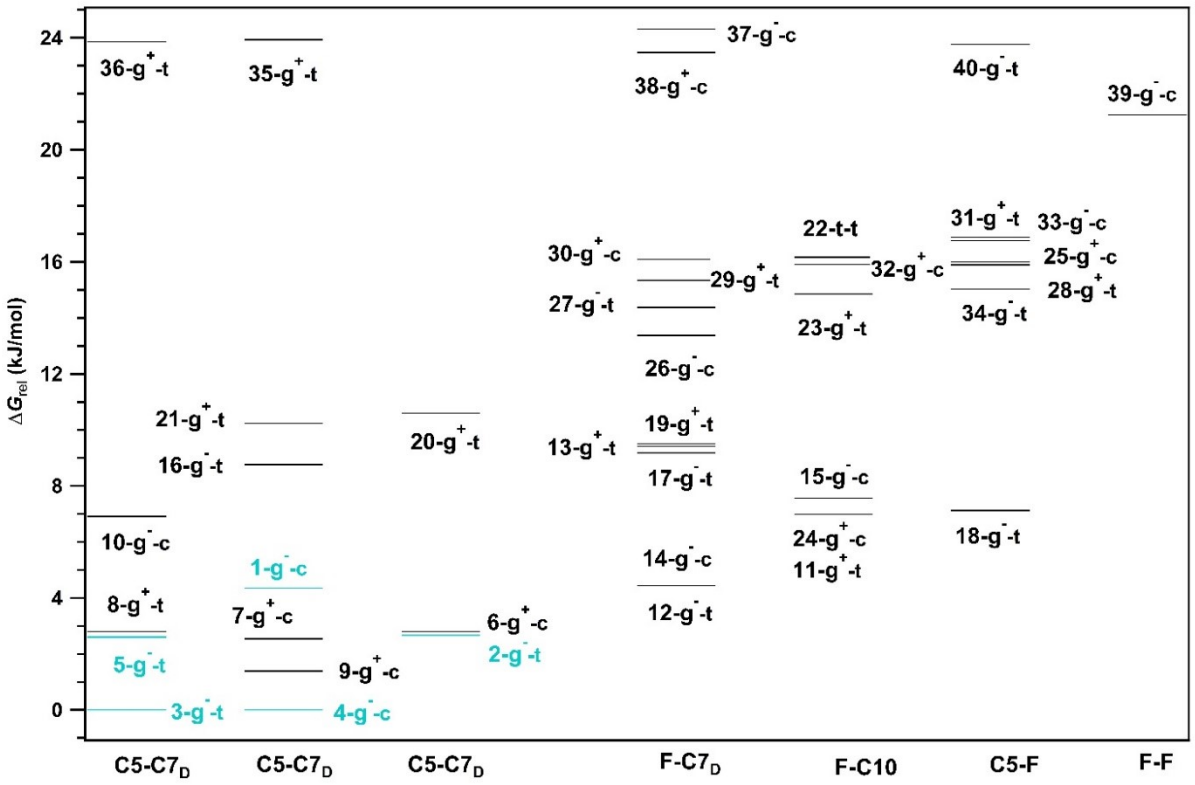


Figure S14. Energy landscape of Boc-Gly-^{D}Pro-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer calculated at 300 K at the B97-D3/def2-TZVPP level of theory. The label C5-C7$_D$ is used thrice merely to show the close-lying energy levels in a distinct manner. See the main text for the nomenclature.

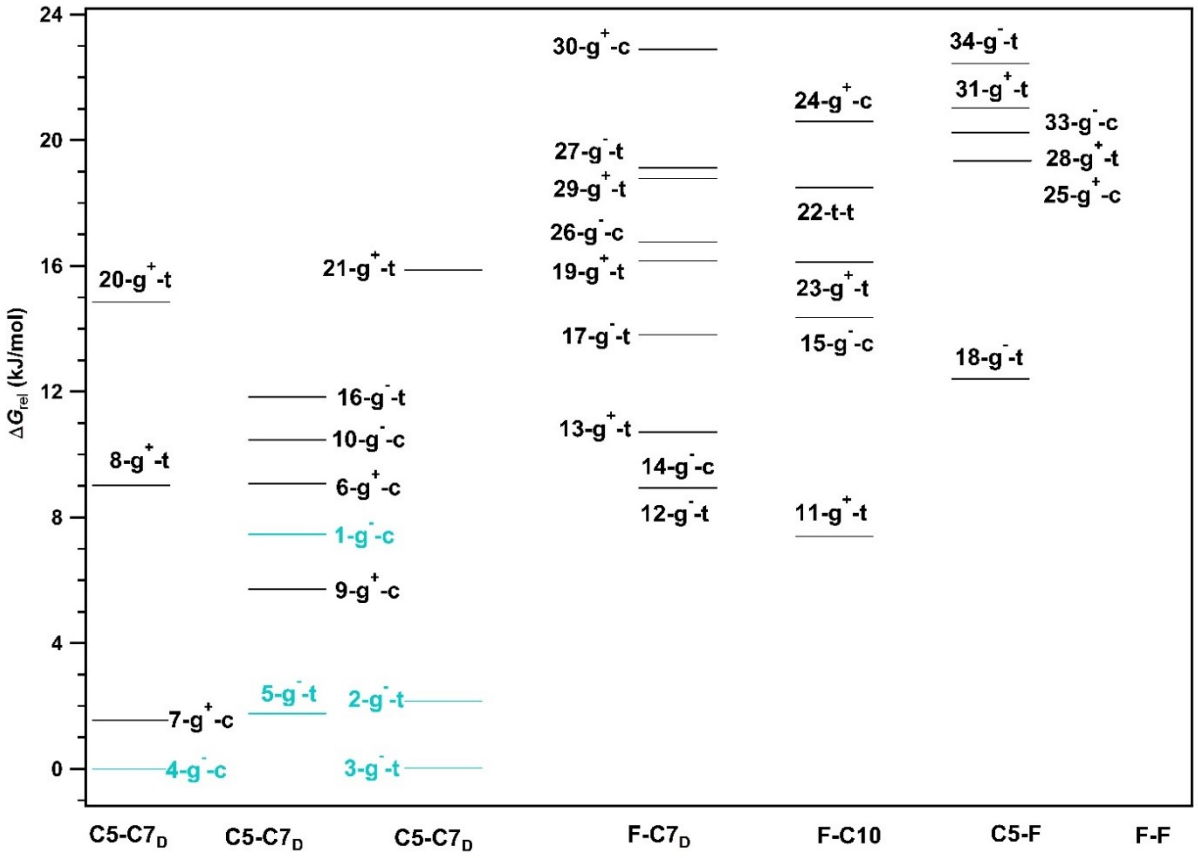


Figure S15. Energetic diagram of Boc-Gly-^{D}Pro-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer at 300 K after calculation at ωB97X-D/def2TZVPP level of theory. The label C5-C7$_D$ is used thrice merely to show the close-lying energy levels in a distinct manner. See the main text for the nomenclature.

4.2.2. Energy landscapes of Pro-Gly peptide at different levels of theory

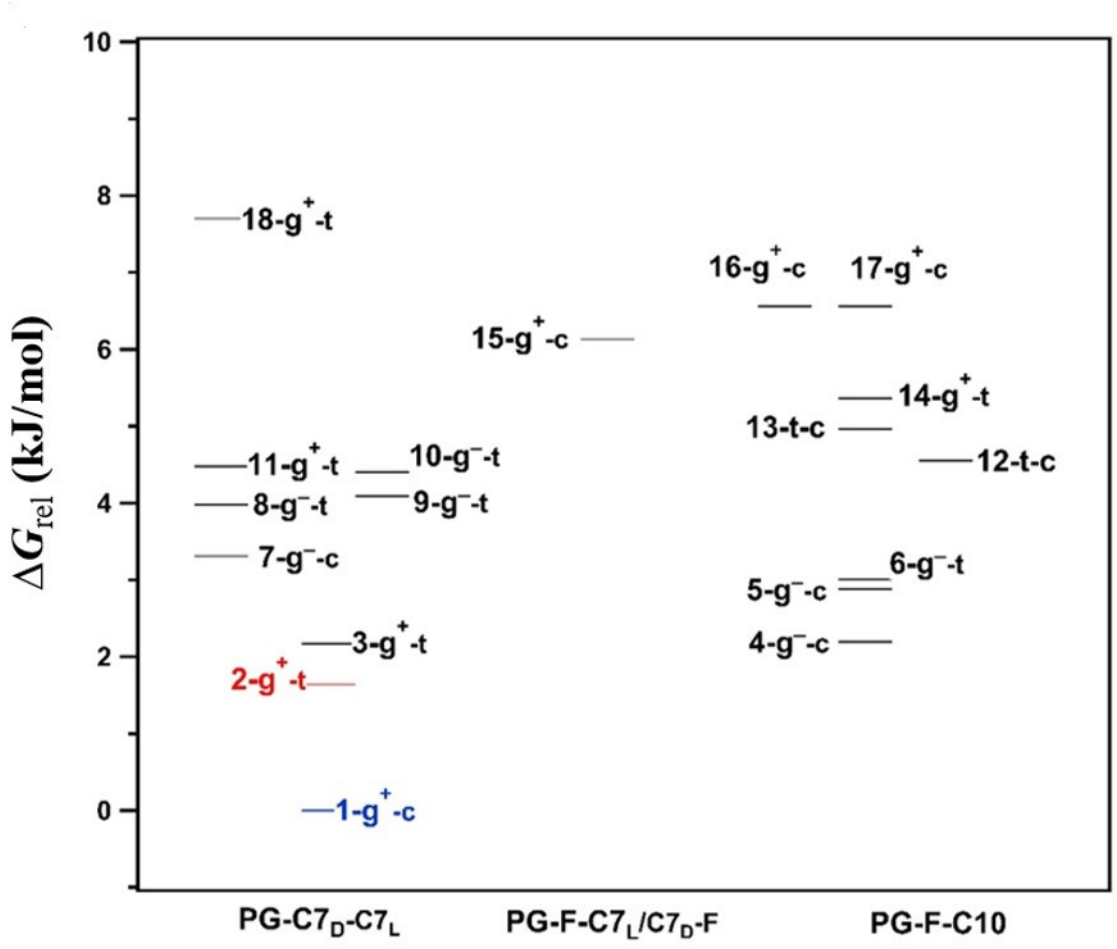


Figure S16. Energy landscape of a few low energy conformers of (a) Boc-^{D}Pro- Gly-NHBn-OMe calculated at 300 K at the M06-2X/6-311++G(2d,2p) level of theory. See the main text for the nomenclature. The color coded conformers are the predicted ones observed in the experiment.

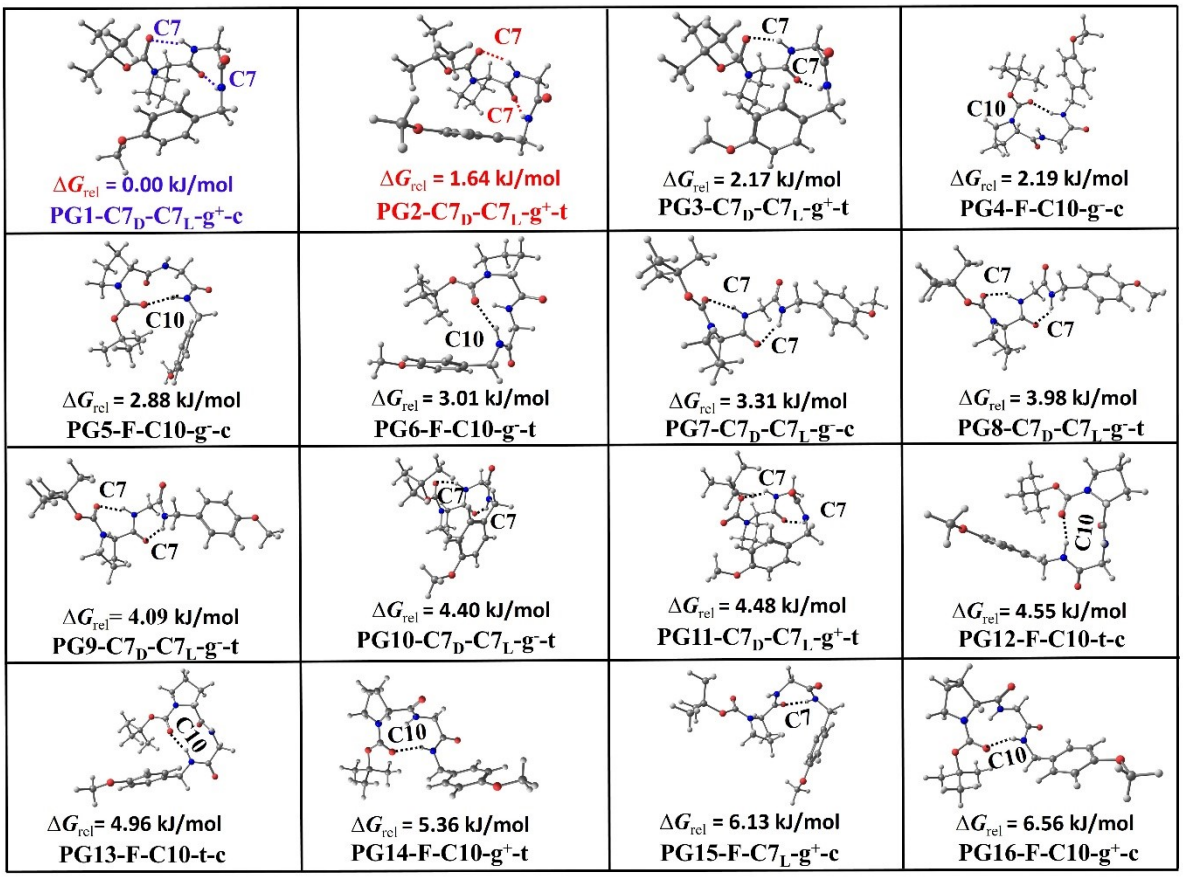


Figure S17. Structures of the 16 lowest energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe optimized at the M06-2X/6-311++G(2d,2p) level of theory. The color coded conformers are the predicted ones observed in the experiment.

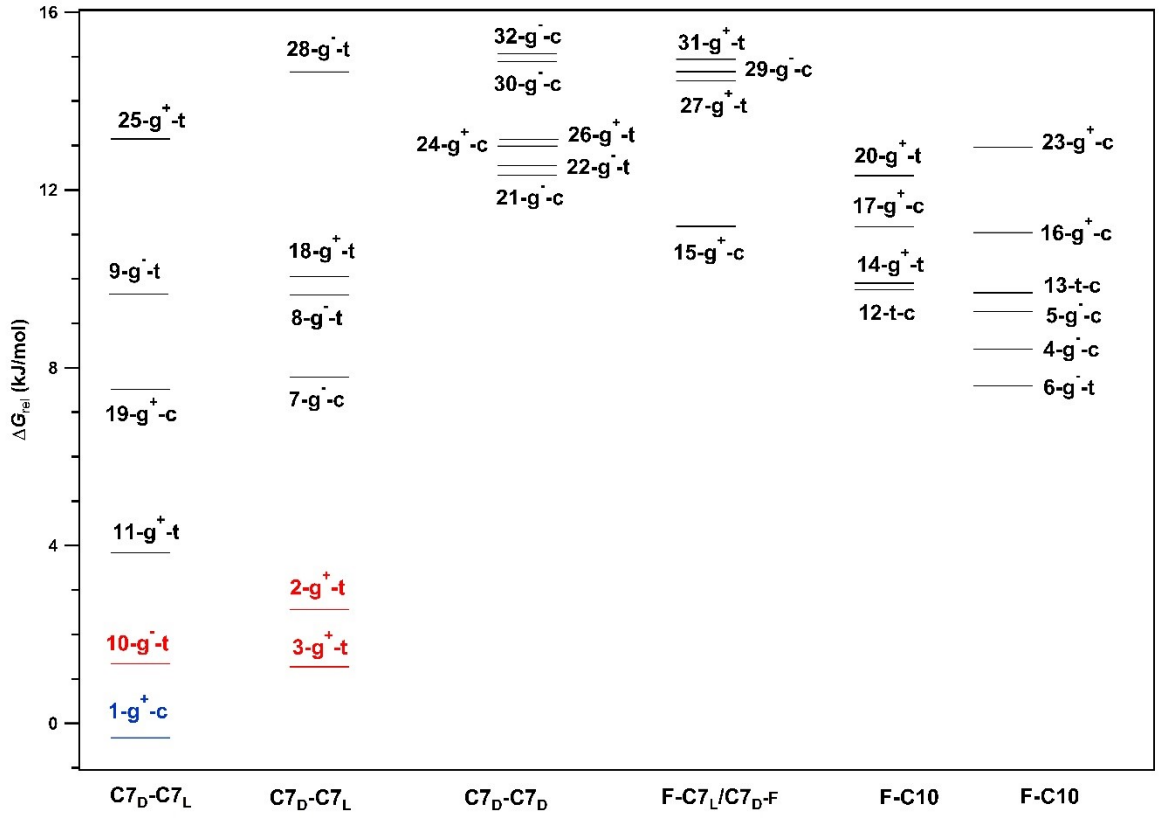


Figure S18. Energy landscape of Boc-^{D}Pro-Gly-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer calculated at 300 K at the B3LYP-D3/def2-TZVPP level of theory. The labels $C7_D$-$C7_L$ and F-C10 are used twice merely to show the close-lying energy levels in a distinct manner. See the main text for the nomenclature.

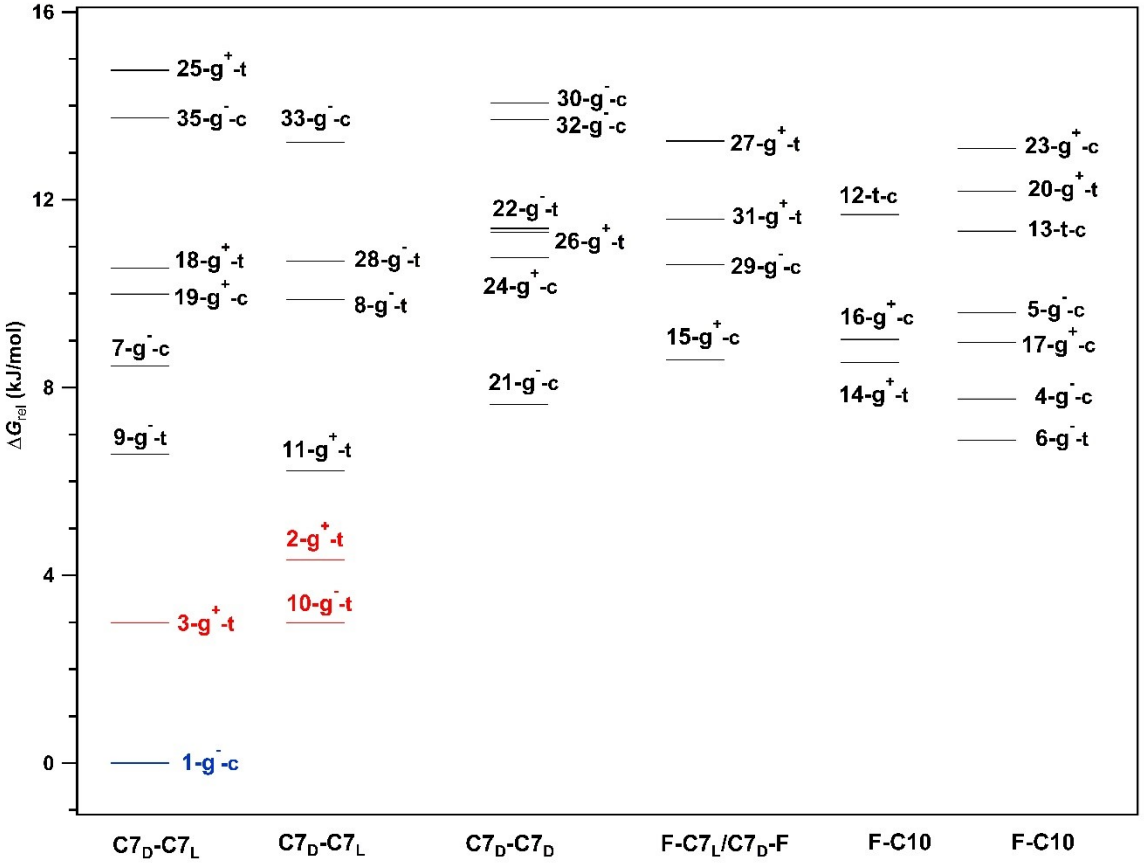


Figure S19. Energy landscape of Boc-^{D}Pro-Gly-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer at 300 K after calculation at B97-D3/def2TZVPP level of theory. The labels $C7_D$-$C7_L$ and F-C10 are used twice merely to show the close-lying energy levels in a distinct manner.

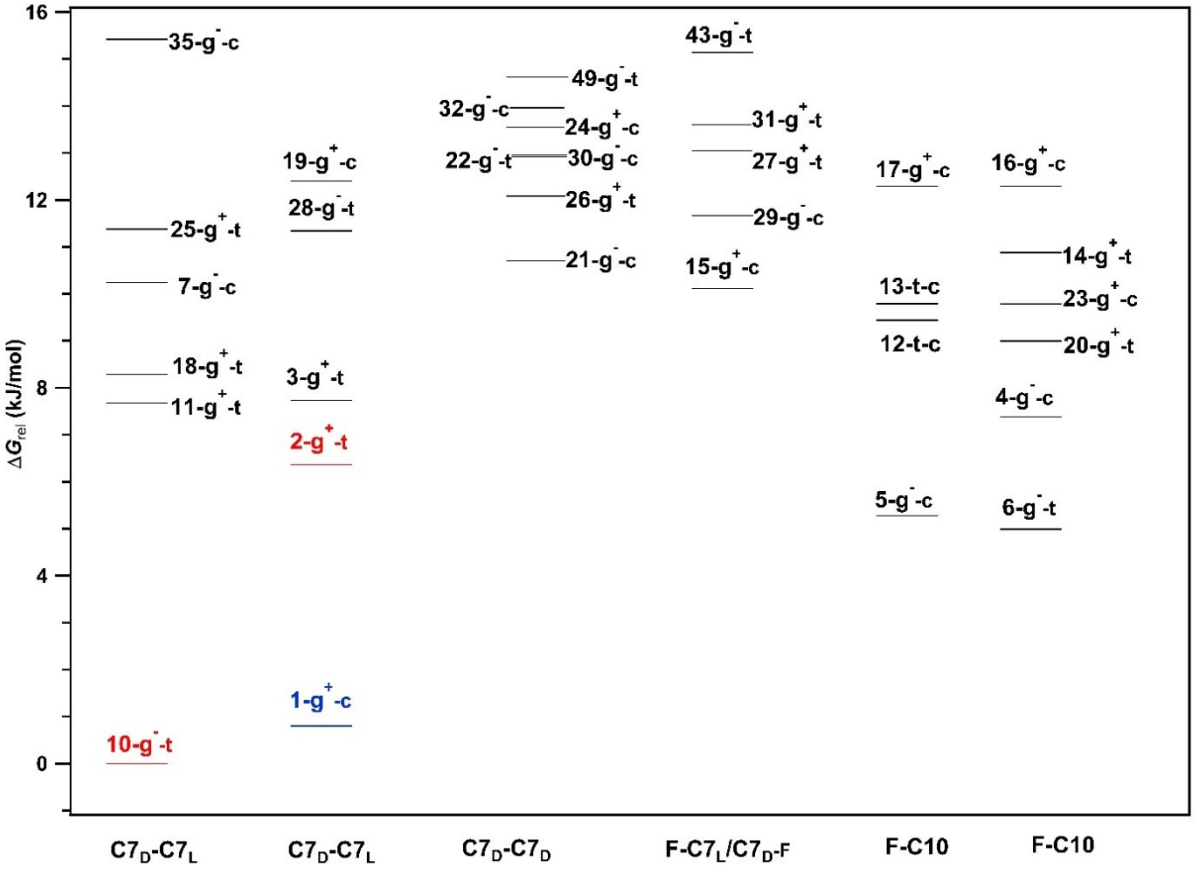


Figure S20. Energy landscape of Boc-^{D}Pro-Gly-NHBn-OMe showing Gibbs free energies (ΔG_{rel}) relative to the most stable conformer calculated at 300 at ωB97X-D/def2TZVPP level of theory. The labels $C7_D$-$C7_L$ and F-C10 are used twice merely to show the close-lying energy levels in a distinct manner.

4.2.3 Comparison of the theoretical IR NH stretching frequencies with those observed in the experiment

In comparing the IR predictions using various functionals, we restrict to the conformers having energies less than 14 kJ/mol for Gly-Pro and 7 kJ/mol for Pro-Gly.

4.2.3.1 Gly-Pro IR NH results and comparison to experiment

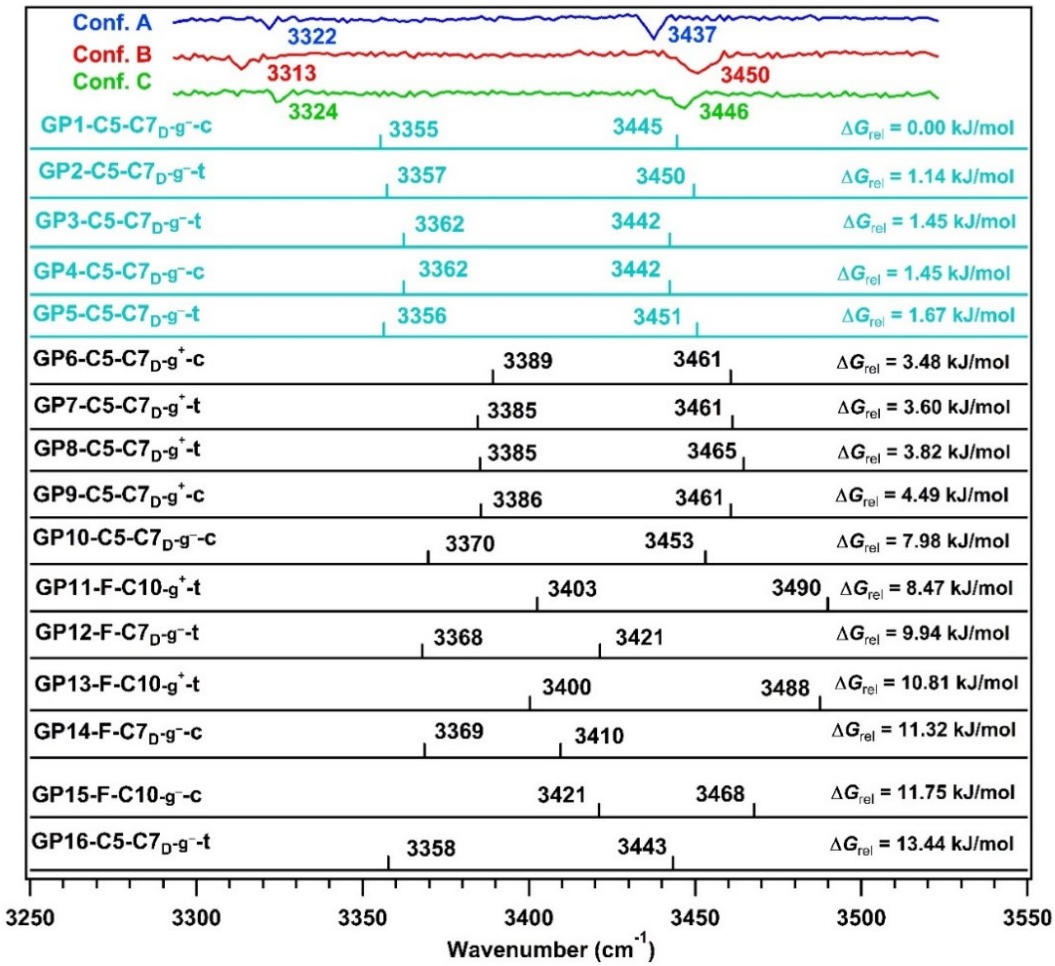


Figure S21. Comparison of the experimental IR spectra with the theoretical IR spectra of the sixteen low energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe calculated at the M06-2X/6-311++G(2d,2p) level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.948 at this particular level of theory. The theoretical stick spectra, which are color coded, are the predicted structures for the three observed conformers in the experiment.

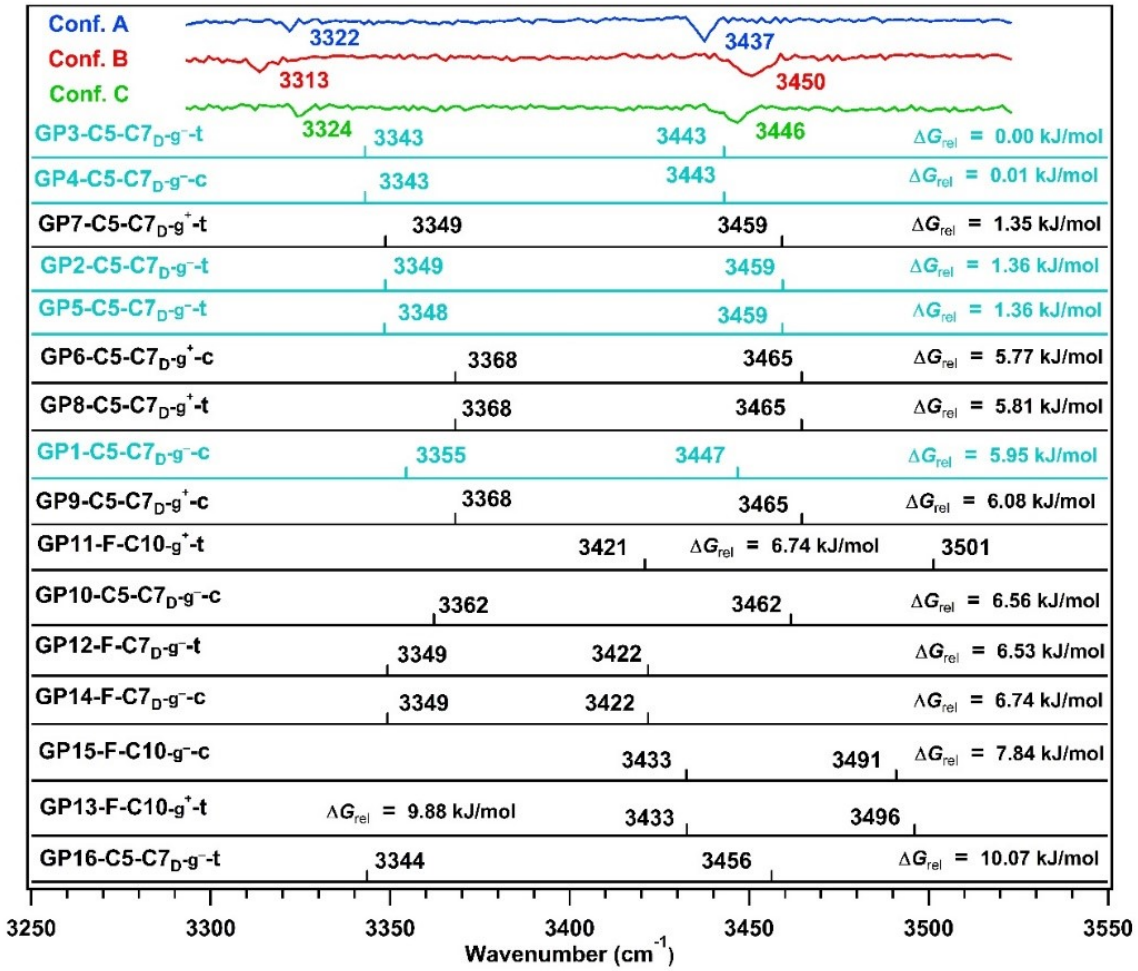


Figure S22. Comparison of the experimental IR spectra with the theoretical IR spectra of the sixteen low energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe calculated at the B3LYP-D3/def2-TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.958 at this particular level of theory.

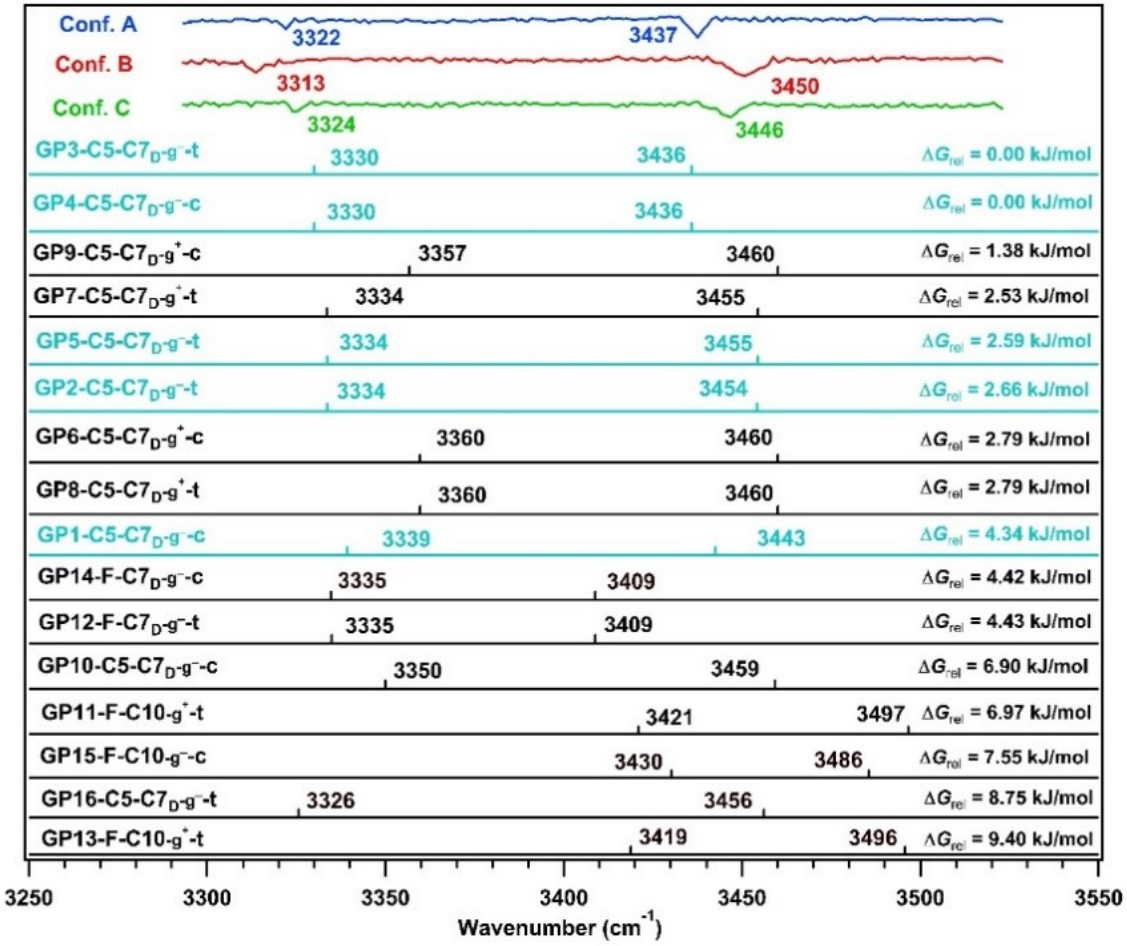


Figure S23. Comparison of the experimental IR spectra with the theoretical IR spectra of the sixteen low energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe calculated at the B97-D3/def2-TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.975 at this particular level of theory.

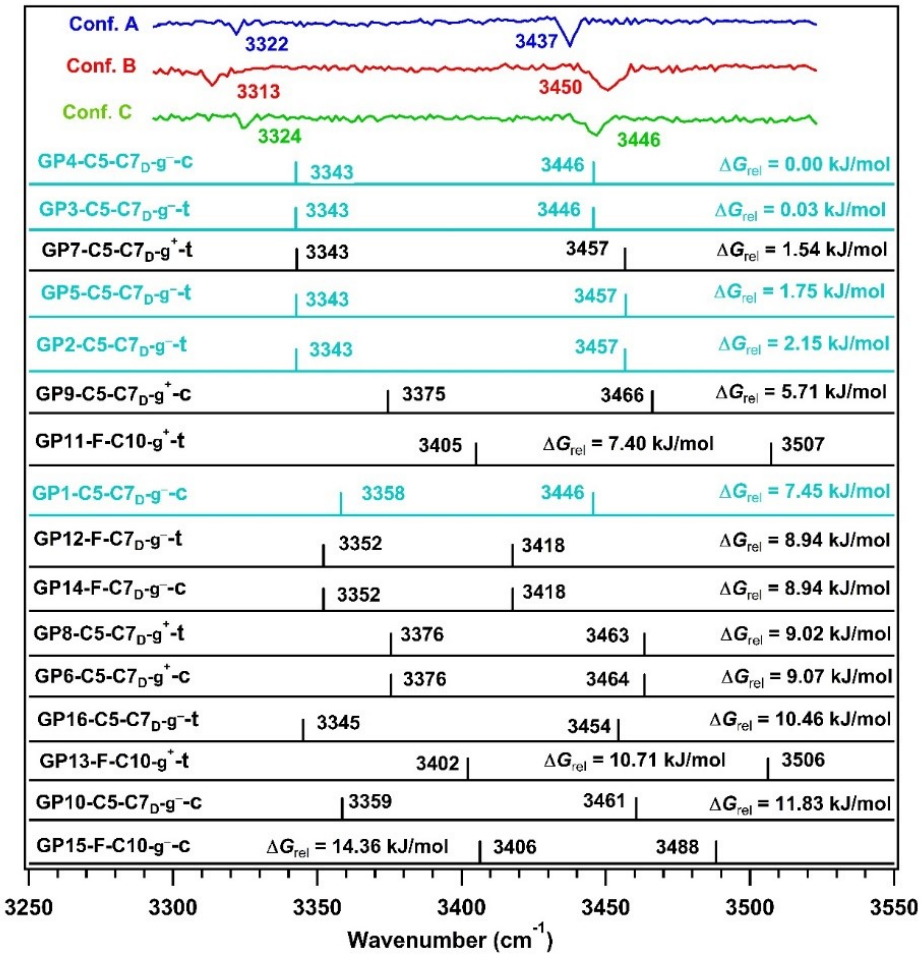


Figure S24. Comparison of the experimental IR spectra with the theoretical IR spectra of the twelve low energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe calculated at the ωB97X-D/def2-TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.945 at this particular level of theory.

4.2.3.2. Gly-Pro unscaled and scaled NH frequencies for M06-2X/6-311++G(2d,2p)

Table S1. Harmonic and scaled NH stretching frequencies of 16 low energy conformers of Boc-Gly-^{D}Pro-NHBn-OMe peptide calculated at the M06-2X/6-311++G(2d,2p) level of theory. Determination of the scaling factor for the N-H stretching frequency has been shown below the table.[a,b]

[a]The scaling factor for the N-H stretching frequencies of both the peptides was derived using the experimental N-H stretching frequency of Z-Gly-OH peptide reported in the literature.[5] The scaling factor for this particular level of theory i.e. M06-2X/6-311++G(2d,2p) was obtained when the experimental NH frequency of Z-Gly-OH was divided by the theoretical harmonic NH frequency calculated at the same level of theory. The reported experimental NH stretching

Conformers	$Glycine_{NH}$ (cm^{-1})		$NHBn_{NH}$ (cm^{-1})	
	Harmonic ν_{N-H}	Scaled ν_{N-H}	Harmonic ν_{N-H}	Scaled ν_{N-H}
GP1-C5-$C7_D$-g^--c	**3633**	**3445**	**3539**	**3355**
GP2-C5-$C7_D$-g^--t	**3637**	**3450**	**3541**	**3357**
GP3-C5-$C7_D$-g^--	**3631**	**3442**	**3546**	**3362**
GP4-C5-$C7_D$-g^--c	**3631**	**3442**	**3547**	**3362**
GP5-C5-$C7_D$-g^--t	**3640**	**3451**	**3541**	**3356**
GP6-C5-$C7_D$-g^+-c	3651	3461	3575	3389
GP7-C5-$C7_D$-g^+-t	3651	3461	3570	3385
GP8-C5-C7D-g^+-t	3655	3465	3571	3385
GP9-C5-$C7_D$-g^+-c	3651	3461	3571	3386
GP10-C5-$C7_D$-g^--c	3643	3453	3555	3370
GP11-F-C10-g^+-t	3681	3490	3589	3403
GP12-F-$C7_D$-g^--t	3609	3421	3553	3368
GP13-F-C10-g^+-t	3679	3488	3587	3400
GP14-F-$C7_D$-g^--c	3597	3410	3553	3369
GP15-F-C10-g^--c	3658	3468	3608	3421
GP16-C5-$C7_D$-g^--t	3632	3443	3542	3358

frequency for the Z-Gly-OH peptide is 3472 cm^{-1} whereas the theoretical harmonic NH

stretching frequency calculated at the M06-2X/6-311++G(2d,2p) level of theory is 3661 cm^{-1}. Thus, the scaling factor at the M06-2X/6-311++G(2d,2p) level of theory is 0.948 (3472/3661 = 0.948).

[b]Experimental Gly_{NH} and $NHBn_{NH}$ frequencies of conformer A of Boc-Gly-^{D}Pro-NHBn-OMe are 3322 and 3437 cm^{-1}, respectively, experimental Gly_{NH} and $NHBn_{NH}$ frequencies of conformer B are 3313 and 3450 cm^{-1}, respectively and experimental Gly_{NH} and $NHBn_{NH}$ frequencies of conformer B are 3324 and 3446 cm^{-1}, respectively. The bold-faced ones in the table are the observed conformers.

4.2.3.3. Pro-Gly IR NH spectra and comparison to experiment

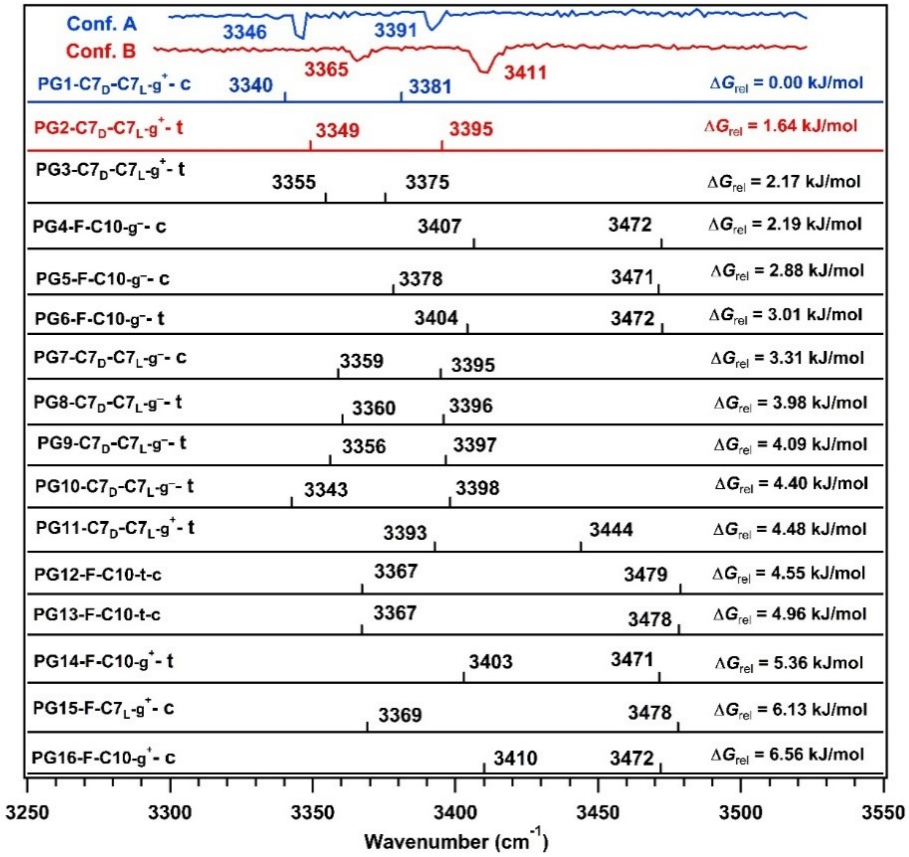


Figure S25. Comparison of the experimental IR spectra with the theoretical IR spectra of the twelve low energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe calculated at the M06-2X/6-311++G(2d,2p) level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.948 at this particular level of theory. The theoretical stick spectra, which are color coded, are the predicted structures for the observed conformers in the experiment.

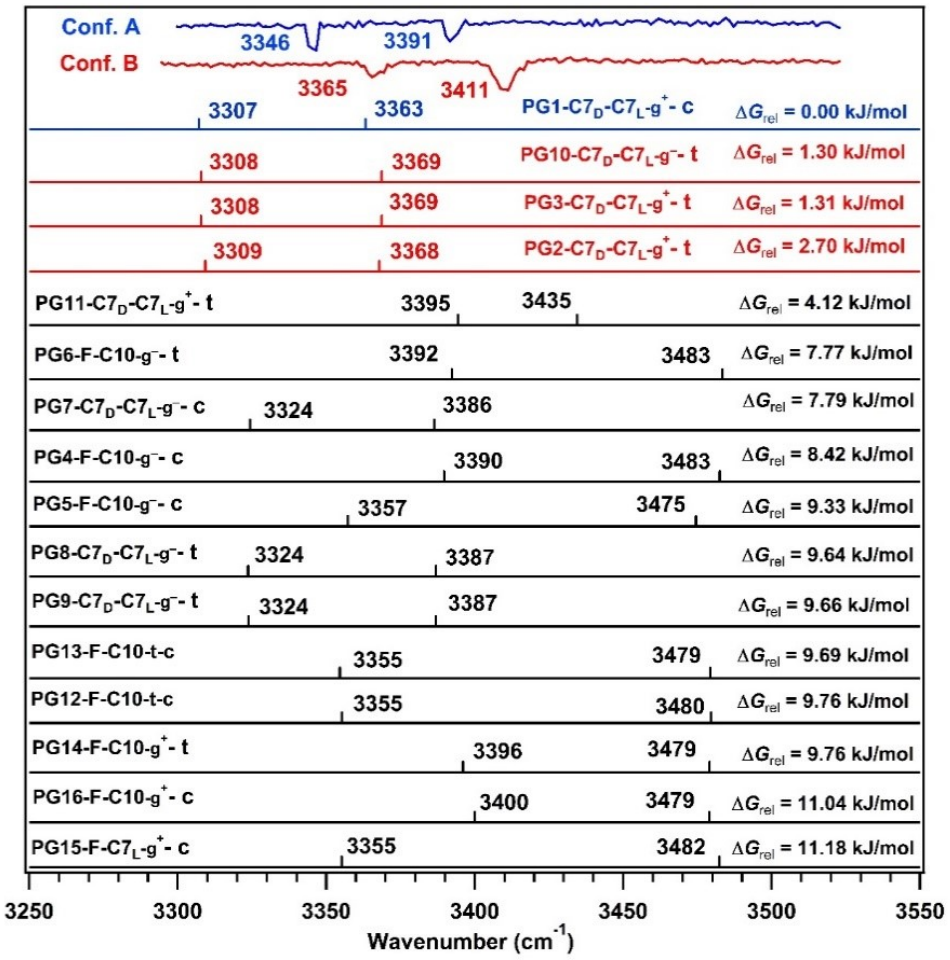


Figure S26. Comparison of the experimental IR spectra with the theoretical IR spectra of the twelve low energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe calculated at the B3LYP-D3/def2TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.958 at this particular level of theory.

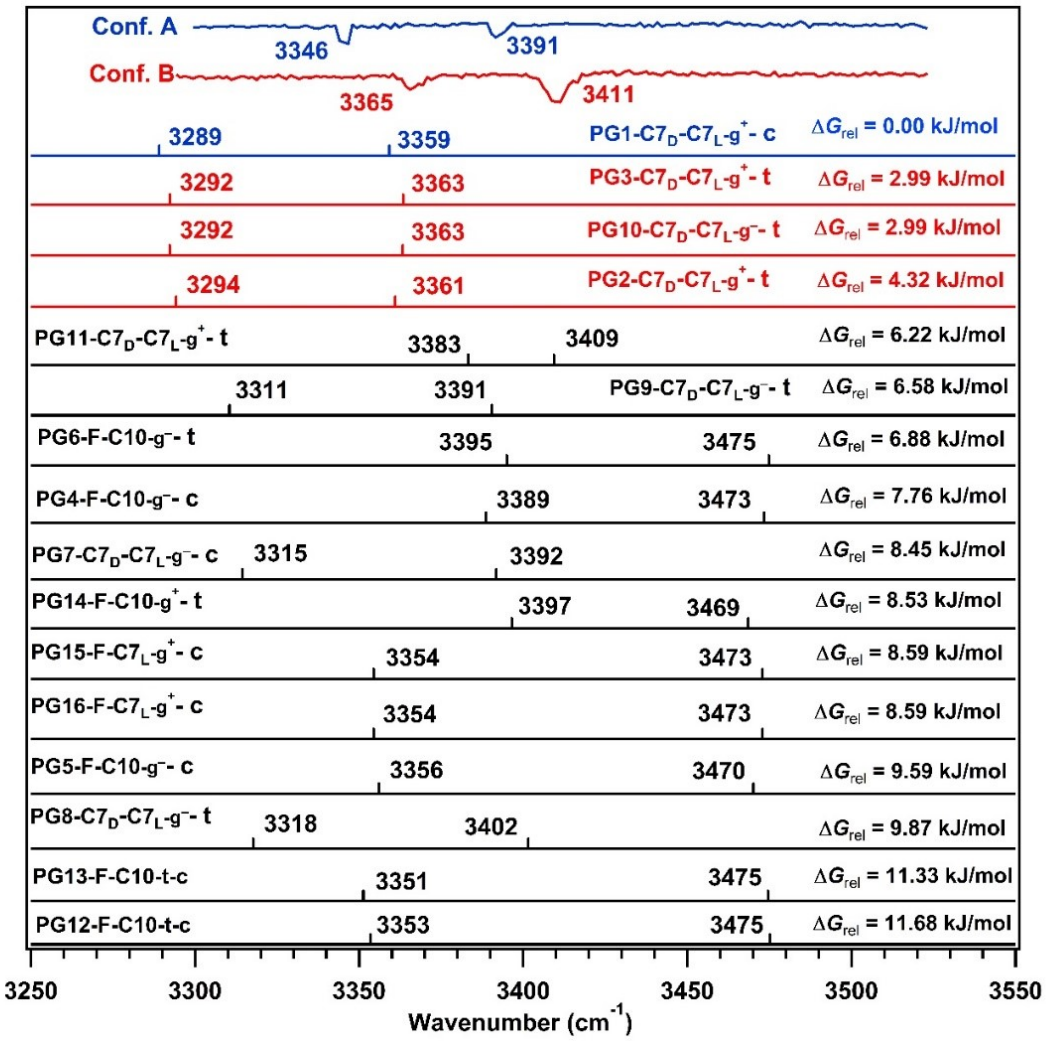


Figure S27. Comparison of the experimental IR spectra with the theoretical IR spectra of the twelve low energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe calculated at the B97-D3/def2TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.975 at this particular level of theory.

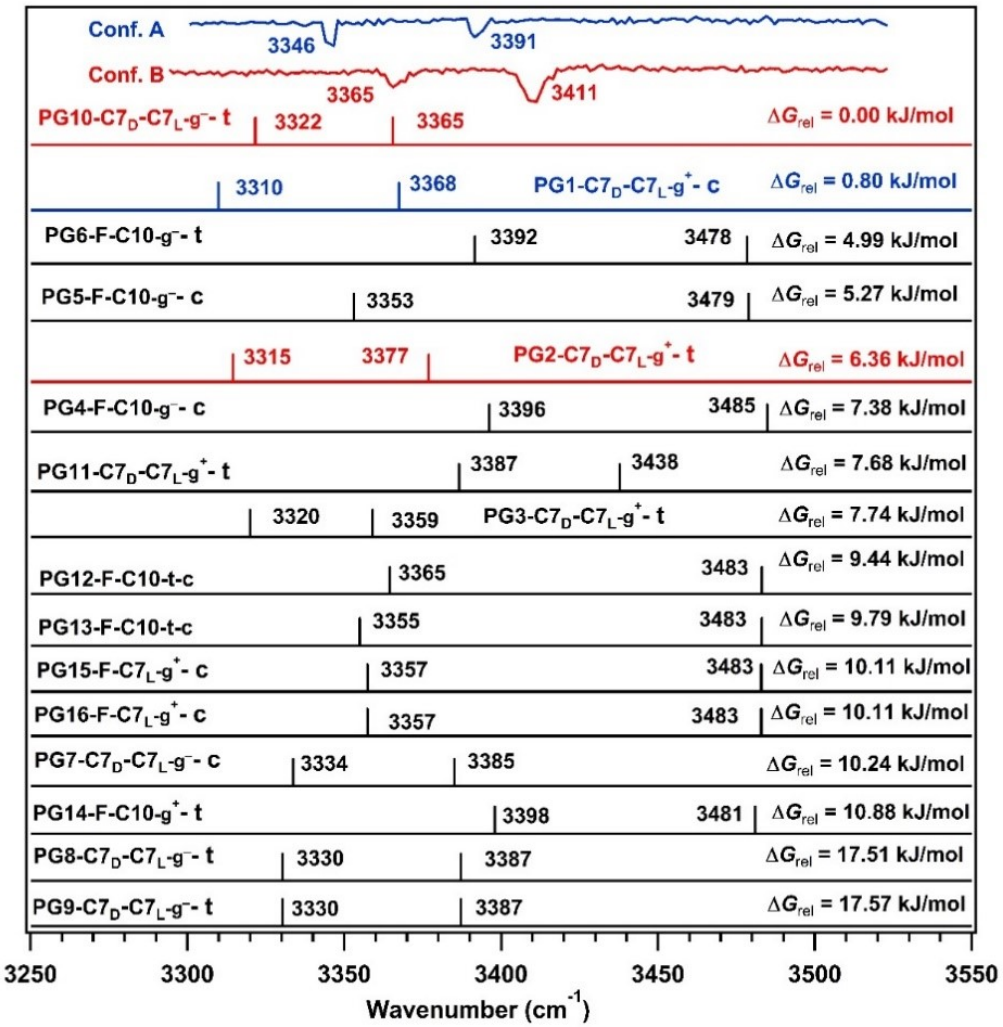


Figure S28. Comparison of the experimental IR spectra with the theoretical IR spectra of the twelve low energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe calculated at the ωB97X-D/def2TZVPP level of theory. Scaling has been done by taking Z-Gly-OH molecule as reference. Scaling factor is 0.945 at this particular level of theory.

4.2.3.4. Pro-Gly unscaled and scaled NH frequencies for M06-2X/6-311++G(2d,2p)

Table S2. Harmonic and scaled NH stretching frequencies of all the 16 low energy conformers of Boc-^{D}Pro-Gly-NHBn-OMe peptide calculated at the M06-2X/6-311++G(2d,2p) level of theory.[a]

Conformers	Glycine$_{NH}$ (cm^{-1})		NHBn$_{NH}$ (cm^{-1})	
	Harmonic ν_{N-H}	Scaled ν_{N-H}	Harmonic ν_{N-H}	Scaled ν_{N-H}
PG1-C7$_D$-C7$_D$-g$^+$-c	**3524**	**3340**	**3567**	**3381**
PG2-C7$_D$-C7$_L$-g$^+$-t	**3533**	**3349**	**3582**	**3395**
PG3-C7$_D$-C7$_L$-g$^+$-t	3538	3355	3560	3375
PG4-F-C10-g$^-$-c	3663	3472	3593	3407
PG5-F-C10-g$^-$-c	3662	3471	3564	3378
PG6-F-C10-g$^-$-t	3663	3472	3591	3404
PG7-C7$_D$-C7$_L$-g$^-$-c	3543	3359	3581	3395
PG8-C7$_D$-C7$_L$-g$^-$-t	3545	3360	3582	3396
PG9-C7$_D$-C7$_L$-g$^-$-t	3540	3356	3583	3397
PG10-C7$_D$-C7$_L$-g$^-$-t	3526	3343	3584	3398
PG11-C7$_D$-C7$_L$-g$^+$-t	3633	3393	3579	3444
PG12-F-C10-t-c	3670	3479	3552	3367
PG13-F-C10-t-c	3669	3478	3552	3367
PG14-F-C10-g$^+$-t	3662	3471	3590	3403
PG15-F-C7$_L$-g$^+$-c	3669	3478	3554	3369
PG16-F-C10-g$^+$-c	3662	3472	3597	3410

[a]The same scaling factor 0.948, which has been used for the Boc-Gly-^{D}Pro-NHBn-OMe peptide at the M06-2X/6-311++G(2d,2p) level of theory, has been used for the Boc-^{D}Pro-Gly-NHBn-OMe peptide at the same level of theory. Experimental Gly_{NH} and $NHBn_{NH}$ frequencies of conformer A of Boc-^{D}Pro-Gly-NHBn-OMe are 3346 and 3391 cm^{-1}, respectively, while experimental Gly_{NH} and $NHBn_{NH}$ frequencies of conformer B are 3365 and 3411 cm^{-1}, respectively. The bold-faced ones in the table are the observed conformers.

4.3. Energy-ordered conformers obtained using solvent calculations

Table S3. A comparison between PCM (chloroform) and gas-phase DFT calculations at the B3LYP- D3/def2-TZVPP level of theory for low-energy conformers of Gly-Pro sequence. The nomenclature is the same as that followed for the gas phase, namely based on the ordering and conformer type observed in M06-2X/6-311++G(2d,2p) in the gas phase. The relative energetics of the conformers change upon solvation, but the nature of the conformers does not change. Hence in both gas and condensed phases, C5-$C7_D$ conformers are favorably formed compared to the C10 orientation.

Conformers	$\Delta \boldsymbol{G}_{rel}$ at 300 K (kJ/mol)	
	PCM	**Gas-Phase**
GP6-C5-$C7_D$-g^+-c	0.00	5.77
GP7-C5-$C7_D$-g^+-t	0.02	1.35
GP5-C5-$C7_D$-g^--t	0.02	1.36
GP4-C5-$C7_D$-g^--c	0.27	0.01
GP8-C5-$C7_D$-g^+-t	0.53	5.81
GP2-C5-$C7_D$-g^--t	0.79	1.36
GP3-C5-$C7_D$-g^--t	0.79	0.00
GP1-C5-$C7_D$-g^--c	0.81	5.95
GP10 C5-$C7_D$-g^--c	0.93	6.56
GP12-F-$C7_D$-g^--t	2.12	6.53
GP9-C5-$C7_D$-g^+-c	5.23	6.08
GP11-F-C10-g^+-t	6.06	6.74
GP13-F-C10-g^+-t	6.06	9.88

Table S4. A comparison between PCM and gas-phase DFT calculations at the B3LYP-D3/def2TZVPP level of theory for low-energy Pro-Gly system. Here we find that the C10 conformers are stabilized more than the C7-C7 conformers; this is in agreement with experimental findings in chloroform solvent. Here also we found that the C10 conformers in the gas phase stabilized in the solvent phase, none of the conformers interconverted into the other due to solvation.

Conformers	$\Delta \boldsymbol{G}_{rel}$ at 300 K (kJ/mol)	
	PCM (chloroform)	**Gas-Phase**
PG6-F-C10-g^--t	0.00	7.77
PG4-F-C10- g^--c	1.94	8.42
PG1-$C7_D$-$C7_L$-g^+-c	2.89	0.00
PG13-F-C10-t-c	3.85	9.69
PG13-F-C10-t-c	4.42	9.76
PG3-$C7_D$-$C7_L$-g^+-t	5.00	1.31
PG10-$C7_D$-$C7_L$-g^--t	5.01	1.30
PG14-F-C10- g^+-t	5.01	9.76
PG5-F-C10- g^--c	6.36	9.33
PG2-$C7_D$-$C7_L$-g^+-t	6.38	2.70
PG16-F-C10- g^+-c	6.47	11.04
PG15-F-$C7_L$-g^+-c	7.18	11.18
PG11-$C7_D$-$C7_L$-g^+-t	8.03	4.12
PG7-$C7_D$-$C7_L$-g^--c	12.31	7.79
PG8-$C7_D$-$C7_L$-g^--t	12.84	9.64
PG9-$C7_D$-$C7_L$-g^--t	12.84	9.66

5. Statistics of the number of the CSD structures containing Gly-Pro and Pro-Gly sequences

A statistics of various non-covalent interactions (C5, C7, and C10) present in the Gly-Pro and Pro-Gly containing peptides was obtained using CCDC conquest 2020 2.0 software. The Cambridge structural database contains numerous structures of organic and inorganic materials.

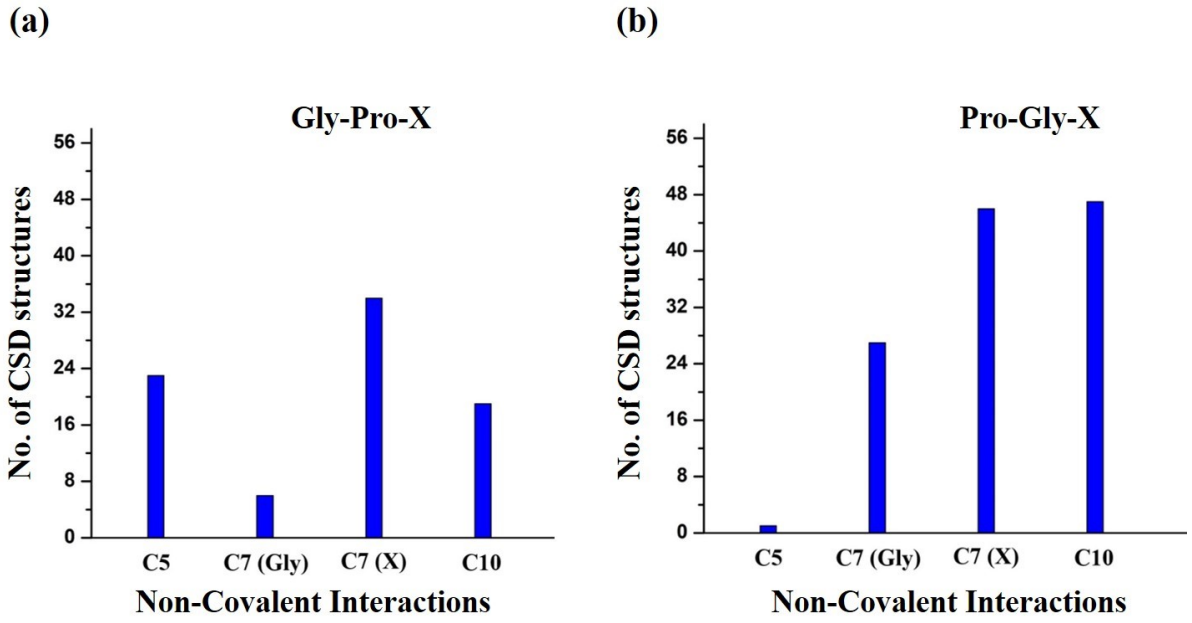


Figure S29. Statistics of the number of CSD structures having different non-covalent interactions in (a) Gly-Pro and (b) Pro-Gly containing peptides.

6. X-ray single crystal structure

A good quality crystal was obtained for the Boc-Gly-^{D}Pro-NHBn-OMe peptide while the quality of the crystal obtained for the Boc-^{D}Pro-Gly-NHBn-OMe peptide was not suitable for the X-ray diffraction. The crystal of Boc-Gly-^{D}Pro-NHBn-OMe was grown in mixed solvent of ethyl acetate and n-hexane. The crystals were grown through slow evaporation of ethyl acetate and n-hexane solvent mixture. X-ray diffraction of the crystal of Boc-Gly-^{D}Pro-NHBn-OMe was performed using APEX(II) DUO CCD diffractometer. The X-ray data were collected at 100 K temperature. The crystal structure of Boc-Gly-^{D}Pro-NHBn-OMe has been provided in Figure 8 in the manuscript while the details of the structure refinement and crystallographic data are provided in Table S6.

The details of the Ramachandran angles and hydrogen bond parameters of the crystal structure of the Boc-Gly-^{D}Pro-NHBn-OMe peptide have been listed in Table 1 provided in the main text. The crystal structure of Boc-Gly-^{D}Pro-NHBn-OMe shows extended PP-II type structure with C5 intramolecular hydrogen bond with the Gly residue which is qualitatively similar to that observed in the gas phase. However, the C7 hydrogen bond interaction in the crystal structure is negligible as the corresponding N-H…O distance (376 pm) is quite large compared to the acceptable hydrogen bond distance. The deviation of the C7 hydrogen bond in the crystal structure could be due to the crystal packing forces through the intermolecular interaction between the neighbouring units.

Table S5. The details of the crystal structure refinement and crystallographic data for Boc-Gly-^{D}Pro-NHBn-OMe

CCDC	2157775
Empirical formula	$C_{20}H_{29}N_3O_5$
M_r	391.46
Crystal size, mm^3	0.12x0.16x0.20
Crystal system	Monoclinic
Space group	P 21
a, Å	11.7946(9)
b, Å	9.3562(8)
c, Å	18.9259(16)
α, °	90
β, °	105.551(2)
γ, °	90
Cell volume, $Å^3$	2012.1(3)
Z ; Z’	4
T, K	100
Radiation type; wavelength Å	MoKα, 0.71073
F_{000}	840
μ, mm^{-1}	0.093
θ range, °	2.234 - 30.170
Reflections collected	11923
Reflections unique	5555
R_{int}	8.35
Parameters	514
wR_2 (all data)	0.1413

The crystal structure was obtained by direct methods using SHELXS-97.[6]

7. References:

1. K.-y. Hung, P. W. R. Harris and M. A. Brimble, *Synlett*, 2009, **2009**, 1233-1236.
2. S. K. Singh, P. R. Joshi, R. A. Shaw, J. G. Hill and A. Das, *Phys. Chem. Chem. Phys.*, 2018, **20**, 18361-18373.
3. S. Kumar, I. Kaul, P. Biswas and A. Das, *J. Phys. Chem. A*, 2011, **115**, 10299-10308.
4. S. Kumar, P. Biswas, I. Kaul and A. Das, *J. Phys. Chem. A*, 2011, **115**, 7461.
5. J. C. Dean, E. G. Buchanan and T. S. Zwier, *J. Am. Chem. Soc.*, 2012, **134**, 17186-17201.
6. G. M. Sheldrick, *Acta Crystallographica Section A*, 1990, **46**, 467-473.