Supporting Information for

Structural analyses of isolated cyclic tetrapeptides with varying amino acid residues

*A. Stamm [a], D. Maué [a], A. Schaly [b], S. Schlicher [b], J. Bartl [b], S. Kubik [b] and M. Gerhards [a] **

[a] TU Kaiserslautern, Fachbereich Chemie & Research Center Optimas, Erwin-Schroedinger-Straße 52, D-67663 Kaiserslautern, Germany.

[b] TU Kaiserslautern, Fachbereich Organische Chemie, Erwin-Schroedinger-Straße 54, D-67663 Kaiserslautern, Germany.

Corresponding Author: Prof. Dr. M. Gerhards *; E-mail: gerhards@chemie.uni-kl.de

Content

cyclo[L-Tyr(Me)-D-Pro-L-Ala-D-Pro] (CP$_{Ala}$)

a) R2PI spectrum of **CP$_{Ala}$**

b) Possible doubly hydrogen-bonded structures and calculated IR spectra of **CP$_{Ala}$** (B3LYP-D3/TZVP)

c) Possible conformers of **CP$_{Ala}$** and calculated IR spectra *without* dispersion corrections (B3LYP/TZVP)

d) Possible conformers and calculated IR spectra of **CP$_{Ala}$** calculated at the ri-MP2/SVP level of theory

e) Possible conformers and calculated IR spectra of **CP$_{Ala}$** including a water solvent shell (B3LYP-D3/TZVP, $\varepsilon = 78.39$)

f) NMR characterization of **CP$_{Ala}$**

g) Mass spectrum of **CP$_{Ala}$**

cyclo[L-Tyr(Me)-D-Pro-L-Glu(Me)-D-Pro] (CP$_{Glu}$)

a) R2PI spectrum of **CP$_{Glu}$**

b) Possible doubly hydrogen-bonded structures and calculated IR spectra of **CP$_{Glu}$** (B3LYP-D3/TZVP)

c) Possible conformers of **CP$_{Glu}$** and calculated IR spectra *without* dispersion corrections (B3LYP/TZVP)

d) Possible conformers and calculated IR spectra of **CP$_{Glu}$** calculated at the ri-MP2/SVP level of theory

e) Possible conformers and calculated IR spectra of **CP$_{Glu}$** including a water solvent shell (B3LYP-D3/TZVP, ε = 78.39)

f) NMR characterization of **CP$_{Glu}$**

g) Mass spectrum of **CP$_{Glu}$**

cyclo[L-Tyr(Me)-D-Pro-L-Ala-D-Pro] (CP_{Ala})

a) R2PI spectrum of **CP_{Ala}**

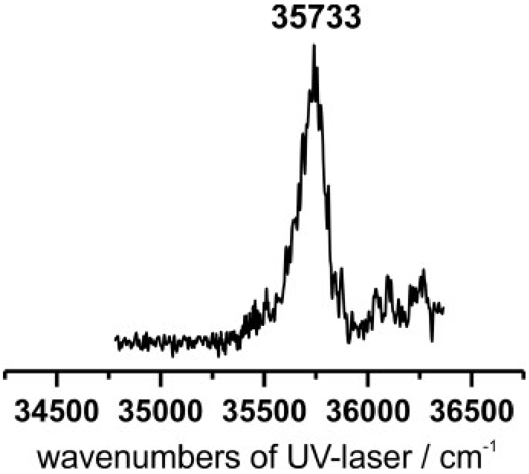


Figure S1. One-color R2PI spectrum of **CP_{Ala}** recorded in the region of 34783-36364 cm^{-1}.

b) Possible doubly hydrogen-bonded structures and calculated IR spectra of **CP$_{Ala}$** (B3LYP-D3/TZVP)

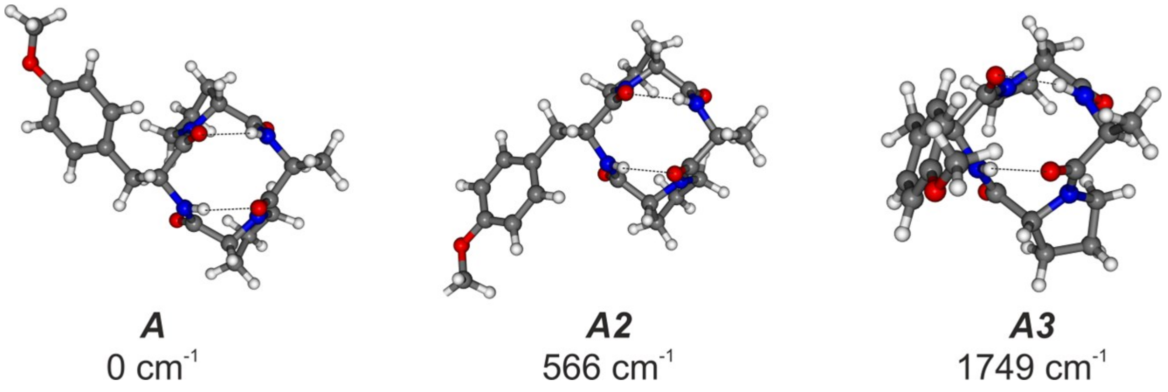


Figure S2. Possible structures of **CP$_{Ala}$** with a doubly hydrogen-bonded binding motif (***A*** type) based on DFT calculations (B3LYP-D3/TZVP). (Hydrogen bond lengths of these structures (given in Å): ***A*** (1.991 and 2.010), ***A2*** (2.00 and 1.991) and ***A3*** (2.009 and 2.031).)

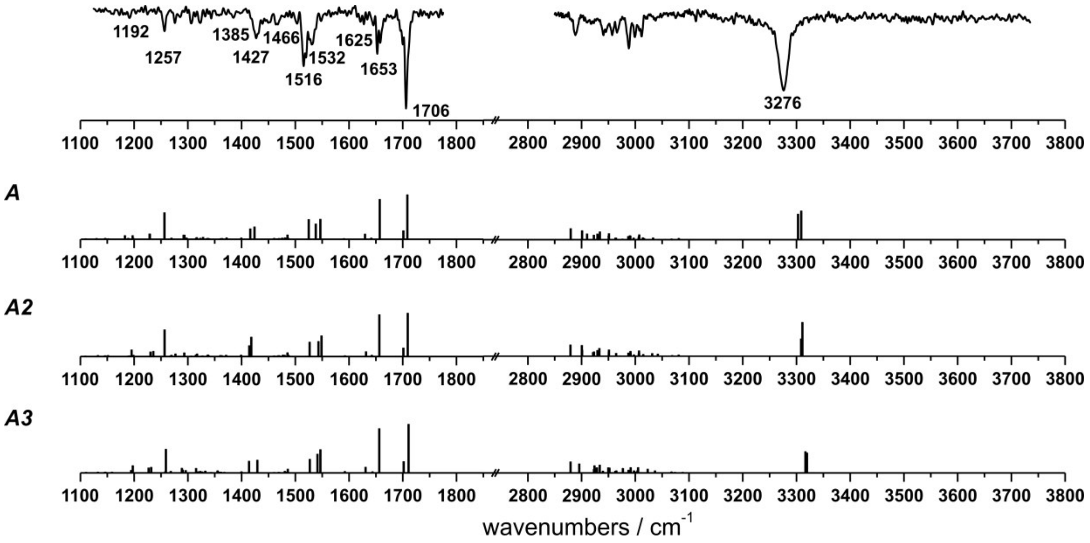
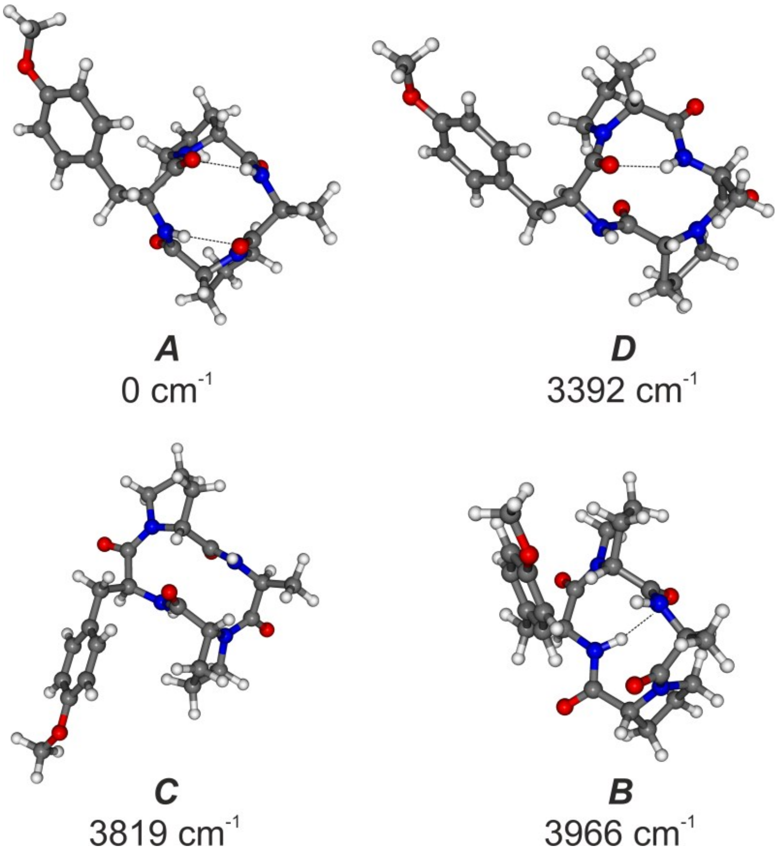


Figure S3. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from DFT calculations (B3LYP-D3/TZVP) of selected conformers of **CP$_{Ala}$** (*cf.* Figure S2). For a better illustration the calculated intensities in the amide A region were doubled. (Details about the used mode-specific scaling factors *cf.* main paper.)

By comparing the harmonically calculated frequencies of ***A***, ***A2*** and ***A3*** of **CP$_{Ala}$** with the experimental IR/R2PI spectrum it can be concluded that all spectra are very similar, but with respect to the NH stretching vibrations conformer ***A*** fits slightly better to the experimental frequencies than ***A2*** and ***A3***. Beside the calculated frequencies also energetic aspects have to be taken into account. Conformer ***A*** is energetically more favored than ***A2*** and ***A3*** which further supports the assignment of conformer ***A***. It should be remarked that the presence of ***A2*** and ***A3*** cannot completely be excluded. Therefore, conformers ***A2*** and ***A3*** may coexist with conformer ***A.***

c) Possible conformers of **CP**$_{\textbf{Ala}}$ and calculated IR spectra *without* dispersion corrections (B3LYP/TZVP)



A
0 cm^{-1}

D
3392 cm^{-1}

C
3819 cm^{-1}

B
3966 cm^{-1}

Figure S4. Possible structures of **CP**$_{\textbf{Ala}}$ based on DFT calculations without dispersion corrections (B3LYP/TZVP). (***A*** represents a structure with two internal hydrogen bonds (2.022 Å and 2.024 Å), ***B*** and ***D*** with one hydrogen bond (2.311 Å and 2.046 Å) and ***C*** without any hydrogen bond.)

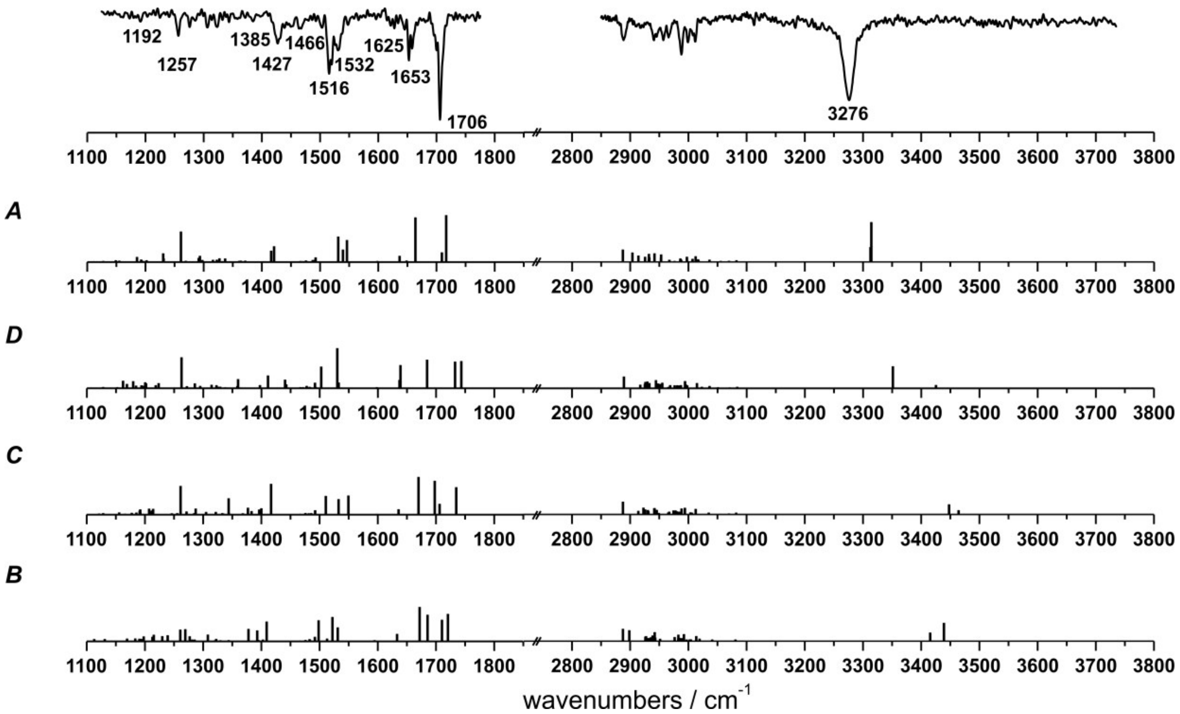


Figure S5. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from DFT calculations without dispersion interactions (B3LYP/TZVP) of selected conformers of $\mathbf{CP_{Ala}}$ (*cf.* Figure S4). Mode-specific scaling factors of 0.9567 for hydrogen-bonded NH stretching vibrations, 0.9613 for non-hydrogen-bonded NH stretching vibrations and 0.99 for the amide I-III region were used. For a better illustration the calculated intensities in the amide A region were doubled.

The comparison of the DFT calculations without dispersion corrections (B3LYP/TZVP) and the experimental IR/R2PI spectrum again yield the assignment of the doubly hydrogen-bonded conformer ***A*** comparable to the calculations including dispersion corrections.

d) Possible conformers and calculated IR spectra of **CP**$_{\mathbf{Ala}}$ calculated at the ri-MP2/SVP level of theory

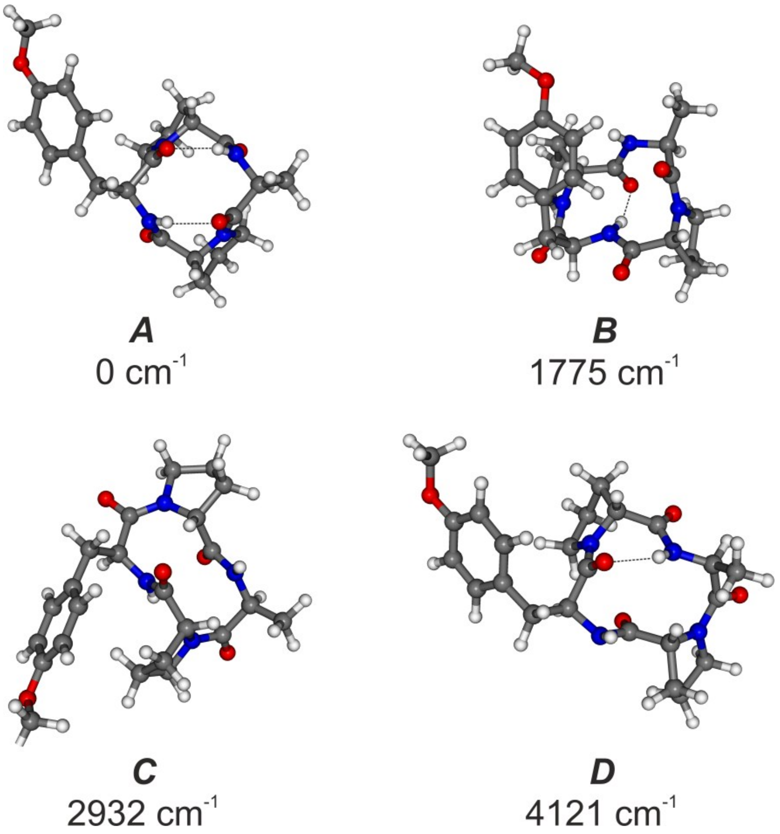


Figure S6. Possible structures of **CP**$_{\mathbf{Ala}}$ based on ri-MP2/SVP calculations. (***A*** represents a structure with two internal hydrogen bonds (1.927 Å and 1.968 Å), ***B*** and ***D*** with one hydrogen bond (2.216 Å and 1.980 Å) and ***C*** without any hydrogen bond.)

All structural binding motifs (***A***, ***B***, ***C*** and ***D***; Figure 3) obtained from the DFT calculations with dispersion corrections (B3LYP-D3/TZVP) were taken as starting geometries for the MP2 optimizations using the approximate resolution of identity (ri) methods together with a SVP basis set. These calculations were realized with Turbomole V7.1 [1,2]. The relative energies are corrected by the zero-point energy and refer to the most stable conformer.

The ri-MP2/SVP optimized structures (Figure S6) basically yield the same geometry as the DFT/B3LYP-D3/TZVP optimizations. Furthermore, a comparison of the calculated frequencies (ri-MP2/SVP) and experimental spectra of **CP**$_{\mathbf{Ala}}$ (Figure S7) again leads to the assignment of the doubly hydrogen-bonded conformer ***A*** comparable to the DFT calculations including and excluding dispersion corrections. Thus an overall agreement between different theoretical methods is obtained. This agreement is a further support that DFT calculations with dispersion corrections (Grimme D3 [3]) and a triple-zeta basis set are a very efficient tool to obtain relative energies of different conformers with reliable frequencies.

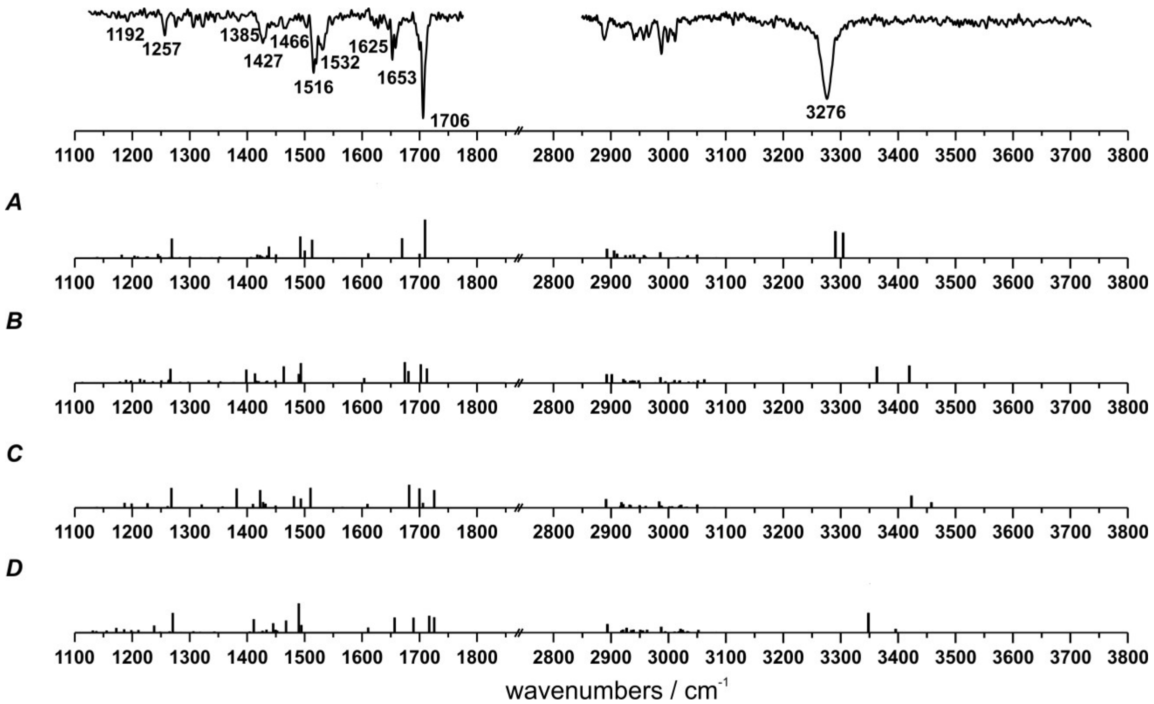


Figure S7. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from ri-MP2/SVP calculations of selected conformers of **CP$_{\mathbf{Ala}}$** (*cf.* Figure S6). Mode-specific scaling factors of 0.9412 for NH stretching vibrations and 0.9492 for the amide I-III region were used (for scaling factors see Ref. 4; the factor of 0.9492 is an average of the CO stretching and NH bending scaling factors of 0.9415 and 0.9568, respectively). For a better illustration the calculated intensities in the amide A region were doubled.

e) Possible conformers and calculated IR spectra of **CP$_{Ala}$** including a water solvent shell (B3LYP-D3/TZVP, ε = 78.39)

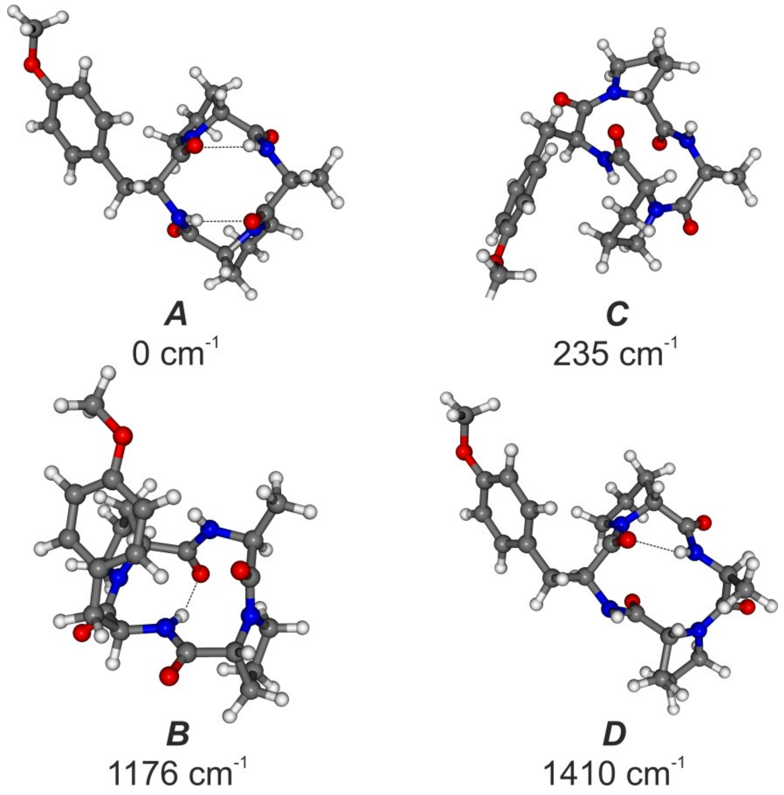


Figure S8. Possible structures of **CP$_{Ala}$** in the energetic order based on DFT calculations (B3LYP-D3/TZVP, ε = 78.39). (***A*** represents a structure with two internal hydrogen bonds (2.010 Å and 2.040 Å), ***B*** and ***D*** with one hydrogen bond (2.258 and 2.034 Å) and ***C*** without any hydrogen bond.)

Solvation effects were taken into account using the polarized continuum model (PCM) implemented in Turbomole's solvation program COSMO [5,6] (**co**nductor-like **s**creening **mo**del). This method determines the electrostatic interaction of a molecule with a solvent which treats the solvent as a continuum with a permittivity ε. These calculations were realized with a combination of Gaussian09 [7] and Turbomole [1,2] (for details see main paper section III) using the B3LYP/TZVP level of theory including dispersion corrections (Grimme D3 [6]). The energy values were corrected by the zero-point energy and refer to the most stable conformer.

For **CP$_{Ala}$** water is used as solvent with a permittivity of 78.39 (B3LYP-D3/TZVP) basically yielding the same results as without solvent effect, *i.e.* structure ***A*** turns out to be the most stable arrangement.

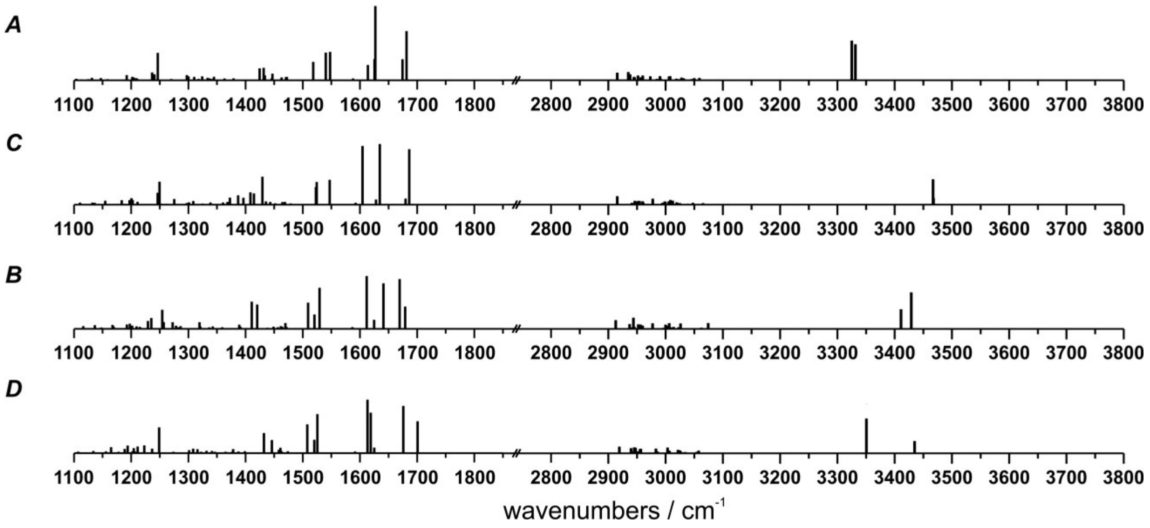


Figure S9. Calculated frequencies obtained from DFT calculations (B3LYP-D3/TZVP, ε = 78.39) of selected conformers of **CP$_{\mathbf{Ala}}$** (*cf.* Figure S8). Mode-specific scaling factors of 0.9563 for hydrogen-bonded NH stretching vibrations, 0.9608 for non-hydrogen-bonded NH stretching vibrations and 0.985 for the amide I-III region were used. For a better illustration the calculated intensities in the amide A region were doubled.

f) NMR characterization of **CP$_{Ala}$**

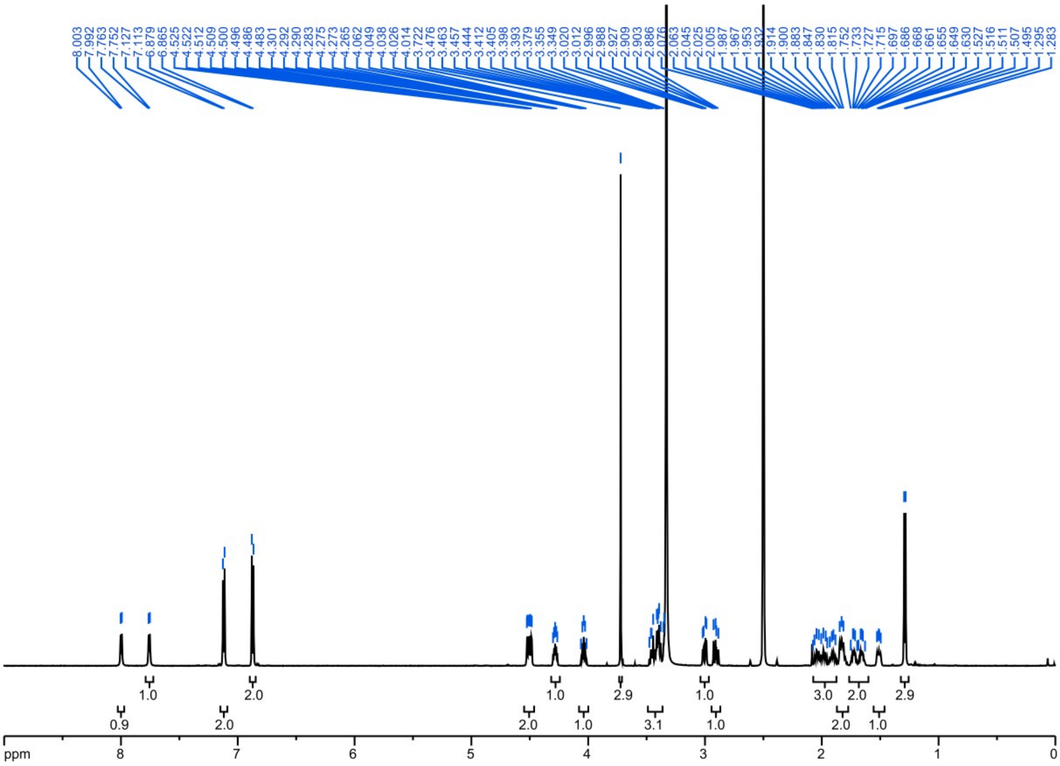


Figure S10: ^{1}H-NMR (600 MHz, DMSO-d$_6$) of **CP$_{Ala}$**.

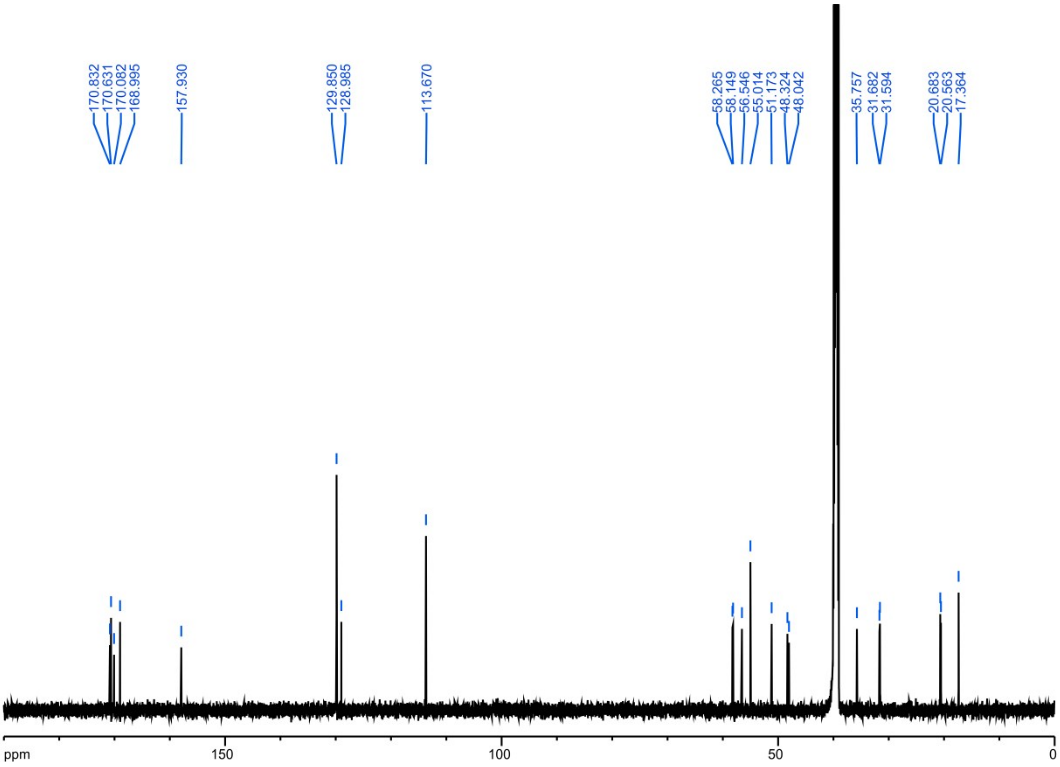


Figure S11: ^{13}C-NMR (151 MHz, DMSO-d$_6$) of **CP$_{Ala}$**.

g) Mass spectrum of $\mathbf{CP}_{\mathbf{Ala}}$

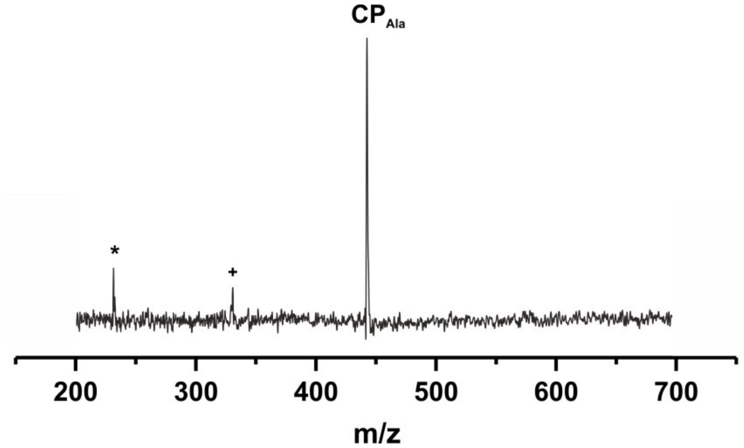


Figure S12: Mass spectrum of $\mathbf{CP}_{\mathbf{Ala}}$ with two fragments (* = 231 m/z and + = 331m/z). The parent mass signal of $\mathbf{CP}_{\mathbf{Ala}}$ obtained at a sample oven temperature of about 120 °C is constant while the IR/R2PI spectra are recorded.

cyclo[L-Tyr(Me)-D-Pro-L-Glu(Me)-D-Pro] (CP$_{Glu}$)

a) R2PI spectrum of **CP$_{Glu}$**

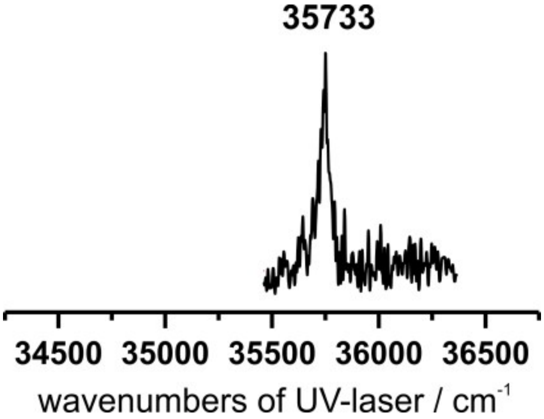
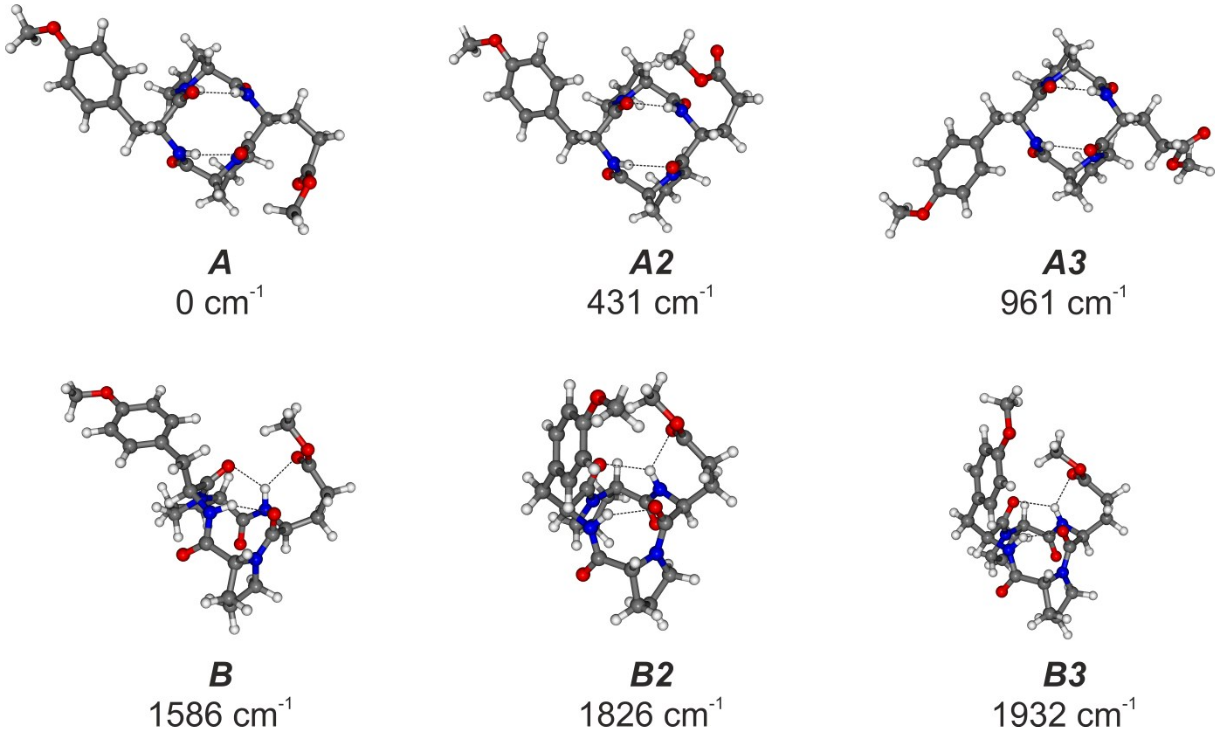


Figure S13. One-color R2PI spectrum of **CP$_{Glu}$** recorded in the region of 35461-36364 cm^{-1}.

b) Possible doubly hydrogen-bonded structures and calculated IR spectra of **CP$_{\mathbf{Glu}}$** (B3LYP-D3/TZVP)



A 0 cm^{-1}

A2 431 cm^{-1}

A3 961 cm^{-1}

B 1586 cm^{-1}

B2 1826 cm^{-1}

B3 1932 cm^{-1}

Figure S14. Possible structures of **CP$_{\mathbf{Glu}}$** with doubly hydrogen-bonded binding motifs (***A*** and ***B*** type) based on DFT calculations including dispersion corrections (B3LYP-D3/TZVP). (Hydrogen bond lengths of these structures (given in Å): ***A*** (1.988 and 2.022), ***A2*** (1.944 and 2.018), ***A3*** (1.995 and 1.997), ***B*** (1.961 and 2.264, 2.108), ***B2*** (2.267 and 1.987, 2.161) and ***B3*** (2.304 and 2.015, 2.122).)

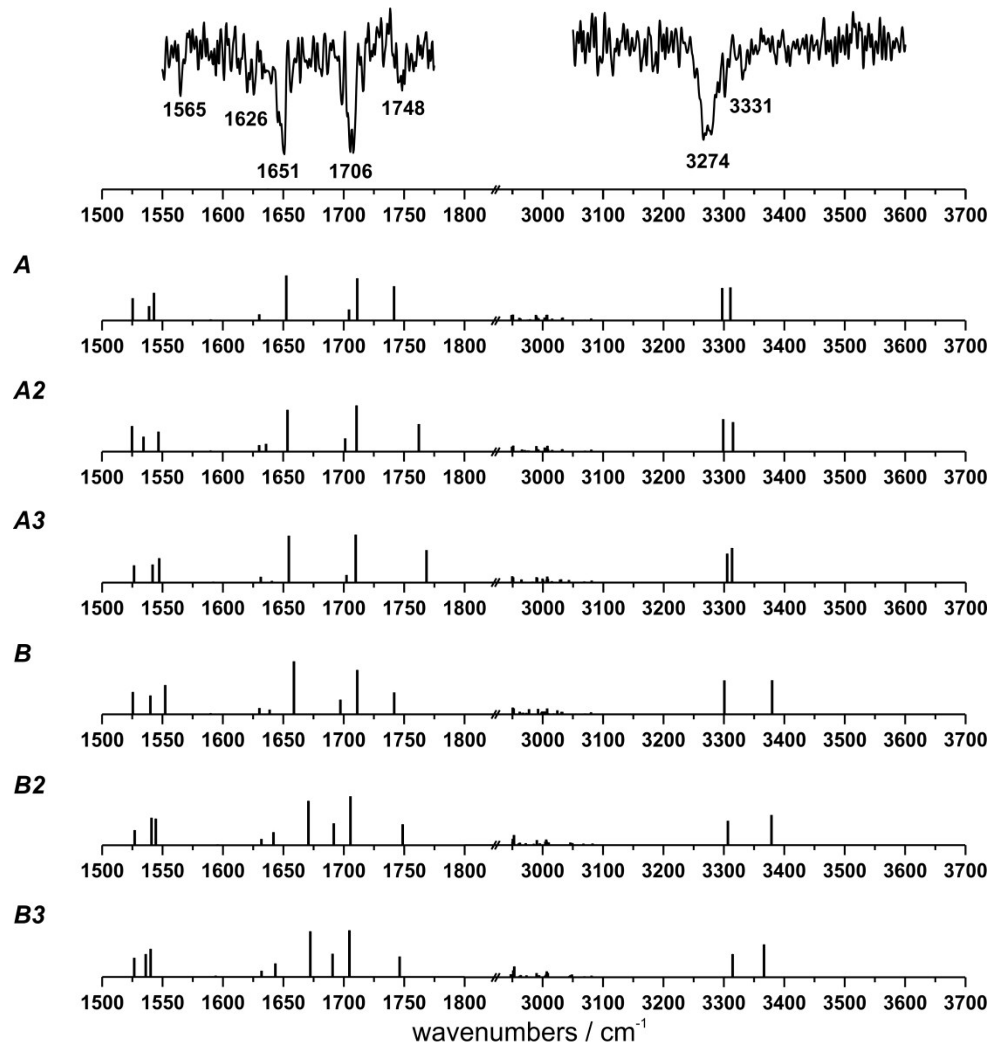
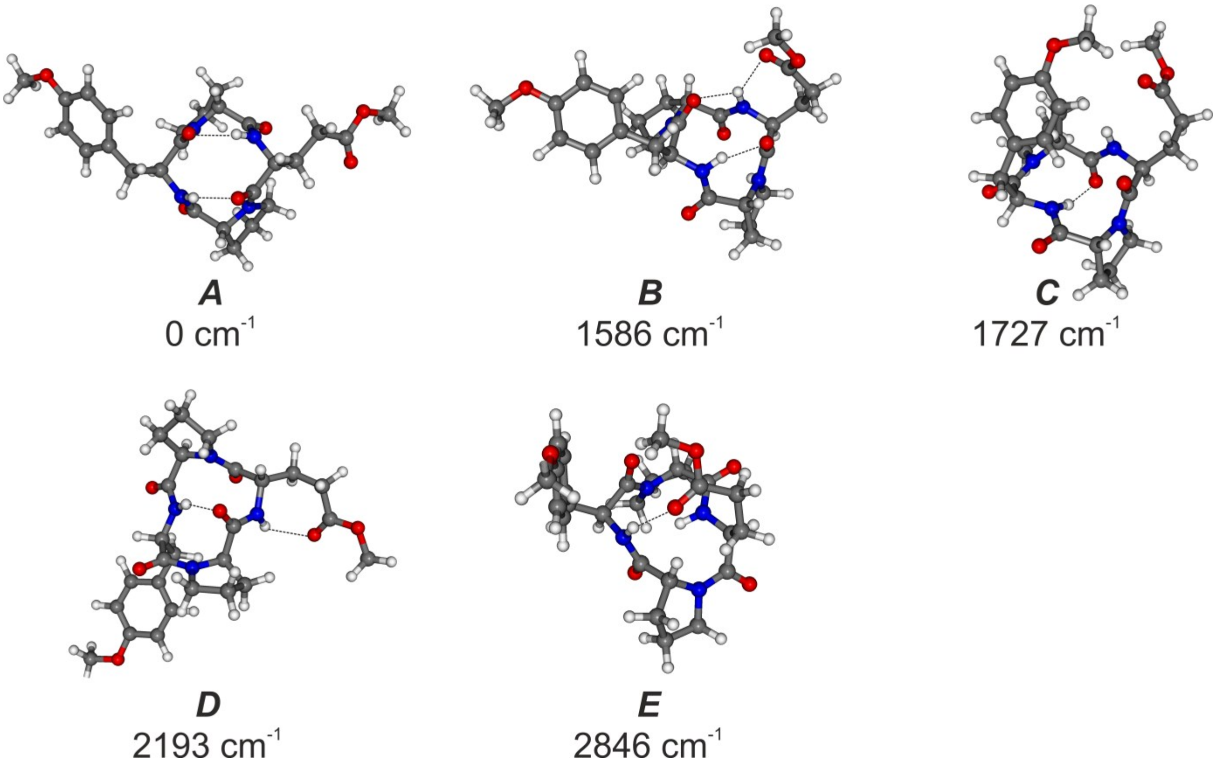


Figure S15. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from DFT calculations (B3LYP-D3/TZVP) of selected conformers of **CP$_{\mathbf{Glu}}$** (*cf.* Figure S14). For a better illustration the calculated intensities in the amide A region were doubled. (Details about the used mode-specific scaling factors *cf.* main paper.)

By comparing the harmonically calculated frequencies of **CP$_{\mathbf{Glu}}$** with the experimentally obtained vibrations ***A2***, ***A3***, ***B2*** and ***B3*** do not explain the IR spectrum as well as the ***A*** and ***B*** structures. In case of conformers ***A2*** and ***A3*** the calculated vibrations are above 1750 cm^{-1} (originating from the carbonyl group of the Glu side chain) and conformers ***B2*** and ***B3*** have predicted frequencies at about 1675 cm^{-1} which are not observed in the experimental spectrum. Moreover these conformers are energetically not favored.

c) Possible conformers of **CP**$_{\mathbf{Glu}}$ and calculated IR spectra *without* dispersion corrections (B3LYP/TZVP)



A
0 cm^{-1}

B
1586 cm^{-1}

C
1727 cm^{-1}

D
2193 cm^{-1}

E
2846 cm^{-1}

Figure S16. Possible structures of **CP**$_{\mathbf{Glu}}$ based on DFT calculations without dispersion corrections (B3LYP/TZVP). ***A*** and ***B*** represent structures with two internal hydrogen bonds (2.019 Å and 2.015 Å for ***A***, 2.006 Å and 2.370 Å for ***B)***, additionally ***B*** forms a further hydrogen bond to the CO side chain of the glutamic acid with a bond-length of 2.147 Å. ***C*** and ***E*** contain one hydrogen bond (bond-lengths 2.294 Å and 2.173 Å) and ***D*** has one internal hydrogen bond (2.142 Å) as well as one bond to the CO side chain of the glutamic acid (bond-length 2.241 Å).

A binding motif of **CP**$_{\mathbf{Glu}}$ without hydrogen bonds (***F*** type) was not found according to these DFT calculations without dispersion corrections within a relative energy of 6000 cm^{-1}.

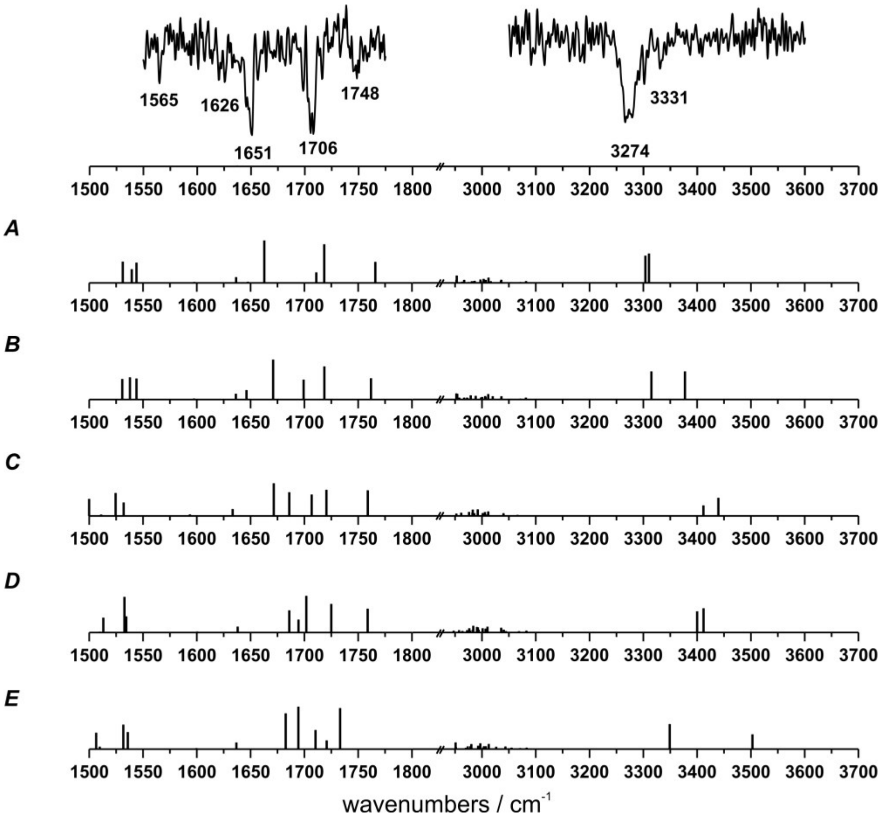


Figure S17. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from DFT calculations without dispersion corrections (B3LYP/TZVP) of selected conformers of **CP$_{\mathbf{Glu}}$** (*cf.* Figure S16). For a better illustration the calculated intensities in the amide A region were doubled. The following mode-specific scaling factors are used: 0.9567 for hydrogen-bonded NH stretching vibrations, 0.9613 for non-hydrogen-bonded NH stretching vibrations and 0.99 for the amide I-III region. For a better illustration the calculated intensities in the amide A region were doubled.

Comparison of the DFT calculations without dispersion corrections (B3LYP/TZVP) and the experimental IR/R2PI spectrum again lead to the assignment of the doubly hydrogen-bonded conformers ***A*** and ***B*** comparable to the calculations including dispersion corrections.

d) Possible conformers and calculated IR spectra of **CP$_{Glu}$** calculated at the ri-MP2/SVP level of theory

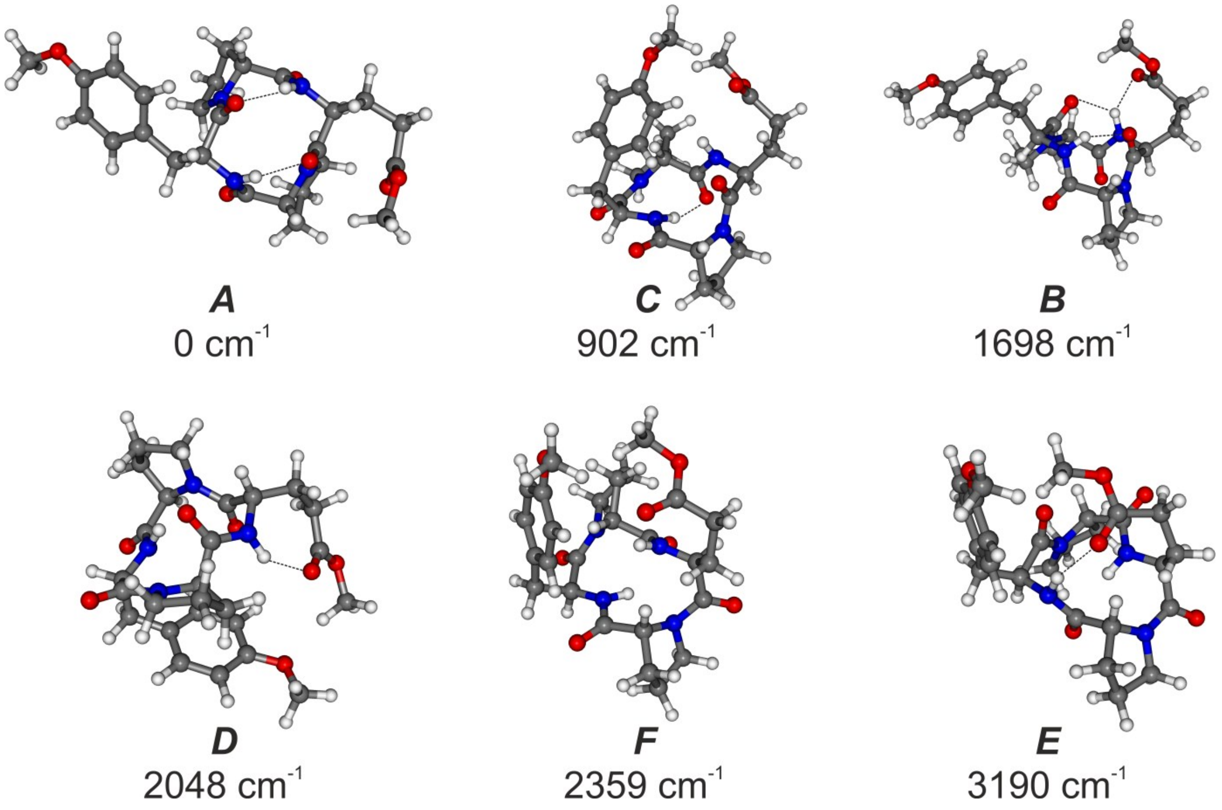


Figure S18. Possible structures of **CP$_{Glu}$** based on ri-MP2/SVP calculations. ***A*** and ***B*** represent structures with two internal hydrogen bonds (1.928 Å and 1.978 Å for ***A***, 1.949 Å and 2.276 Å for ***B)***, additionally ***B*** forms a further hydrogen bond to the CO side chain of the glutamic acid with a bond-length of 2.003 Å. ***C*** and ***E*** contain one hydrogen bond (bond-lengths 2.253 Å and 2.092 Å) and ***D*** has one internal hydrogen bond (1.984 Å) as well as one bond to the CO side chain of the glutamic acid (bond-length 1.911 Å).

All structural binding motifs (***A***, ***B***, ***C***, ***D***, ***E*** and ***F***; Figure 5) obtained from the DFT calculations with dispersion corrections (B3LYP-D3/TZVP) were taken as starting geometries for the MP2 optimizations using the approximate resolution of identity (ri) methods together with a SVP basis set. These calculations were realized with Turbomole V7.1 [1,2]. The relative energies are corrected by the zero-point energy and refer to the most stable conformer.

The ri-MP2/SVP optimized structures (Figure S18) basically yield the same geometry as the DFT/B3LYP-D3/TZVP optimizations only the relative energies change. Furthermore, a comparison of the calculated frequencies (ri-MP2/SVP) and experimental spectra of **CP$_{Glu}$** (Figure S19) again lead to the assignment of the doubly hydrogen-bonded conformers ***A*** and ***B***. Both conformers are stabilized by two internal hydrogen bonds (within the cyclic peptide ring) but conformer ***B*** additionally contains a hydrogen bond between the carbonyl group of the Glu side chain and the NH$_{Glu}$ group comparable to the DFT calculations including and excluding dispersion corrections. It should be remarked that in case of the DFT-D3 calculations the assigned conformers ***A*** and ***B*** are the most stable structures whereas in case of the ri-MP2 method structure ***C*** is more stable than conformer ***B*** although the frequencies still indicate that this structural arrangement has to be assigned to the IR spectrum (similar to the DFT calculation). Like in the case of **CP$_{Ala}$** this agreement is a further support that DFT calculations with dispersion corrections (Grimme D3 [3]) and a triple-zeta basis set are a very efficient tool to obtain relative energies of different conformers with reliable frequencies.

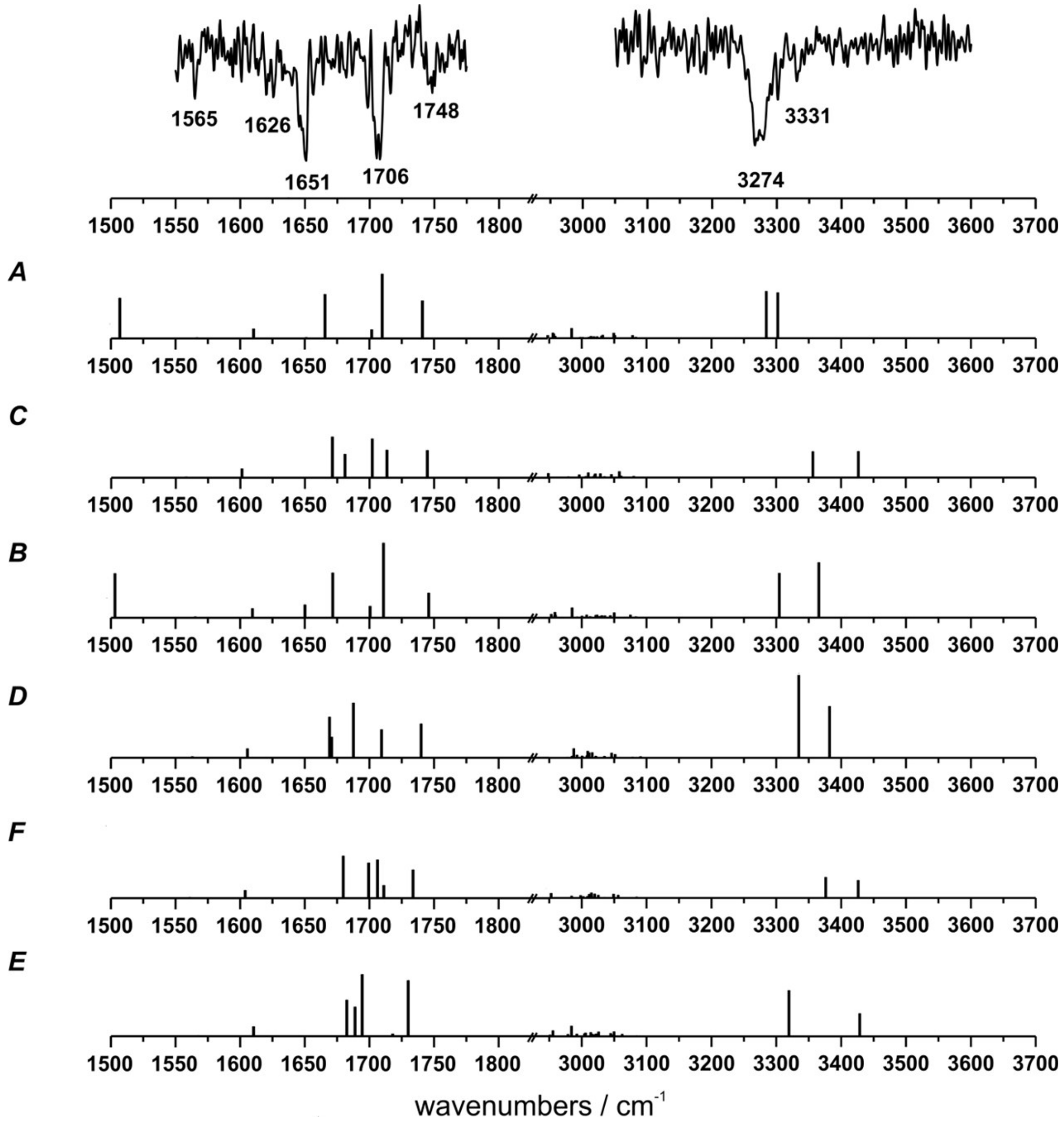
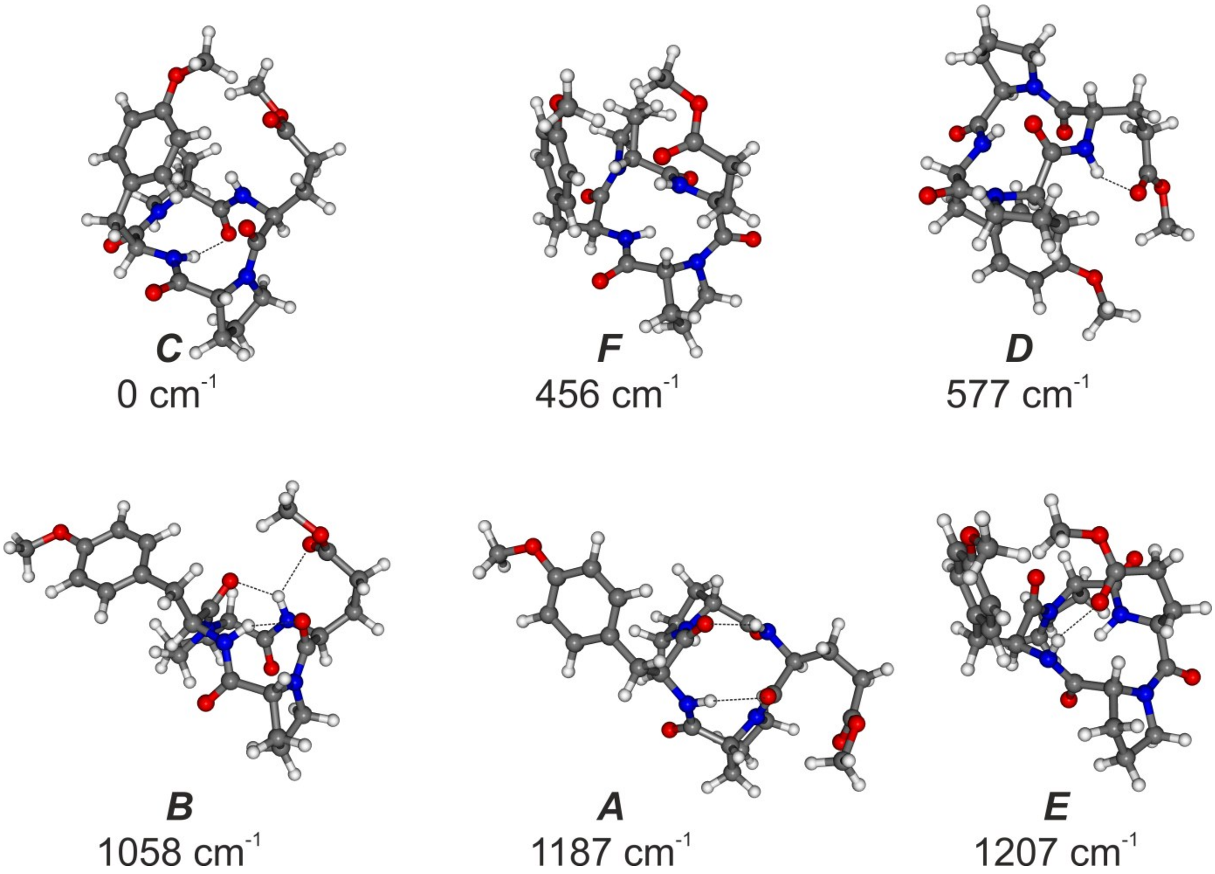


Figure S19. Comparison between the experimental IR/R2PI spectrum and calculated frequencies obtained from ri-MP2/SVP calculations of selected conformers of **CP$_{Glu}$** (*cf.* Figure S18). Mode-specific scaling factors of 0.9412 for NH stretching vibrations and 0.9492 for the amide I-III region were used (for scaling factors see Ref. 4; the factor of 0.9492 is an average of the CO stretching and NH bending scaling factors of 0.9415 and 0.9568, respectively). For a better illustration the calculated intensities in the amide A region were doubled.

e) Possible conformers and calculated IR spectra of $\mathbf{CP}_{\mathrm{Glu}}$ including a water solvent shell (B3LYP-D3/TZVP, ε = 78.39)


C 0 cm^{-1} ***F*** 456 cm^{-1} ***D*** 577 cm^{-1}

B 1058 cm^{-1} ***A*** 1187 cm^{-1} ***E*** 1207 cm^{-1}

Figure S20. Possible structures of $\mathbf{CP}_{\mathrm{Glu}}$ in the energetic order based on DFT calculations (B3LYP-D3/TZVP, ε = 78.39). ***A*** and ***B*** represent structures with two internal hydrogen bonds (2.010 Å and 2.037 Å for ***A***, 1.968 Å and 2.106 Å for ***B)***, additionally ***B*** forms a further hydrogen bond to the CO side chain of the glutamic acid with a bond-length of 2.263 Å. ***C*** and ***E*** contain one hydrogen bond (bond-lengths 2.319 Å and 2.092 Å) and ***D*** has one internal hydrogen bond (2.009 Å) as well as one bond to the CO side chain of the glutamic acid (bond-length 1.895 Å).

Similar to $\mathbf{CP}_{\mathrm{Ala}}$ solvation effects for $\mathbf{CP}_{\mathrm{Glu}}$ were also taken into account by using the PCM model implemented in Turbomole's solvation program COSMO [5,6] (for details about the calculation procedure see section $\mathbf{CP}_{\mathrm{Ala}}$/e)). Water is used as solvent with a permittivity of 78.39 (B3LYP-D3/TZVP) yielding a structure with one internal hydrogen-bond as most stable arrangement. This is in contrast to the results of $\mathbf{CP}_{\mathrm{Ala}}$ for which the most stable arrangement of the non-solvated cyclic peptide remains the most stable one upon solvation. This difference results from the additional carbonyl group in $\mathbf{CP}_{\mathrm{Glu}}$ (compared to $\mathbf{CP}_{\mathrm{Ala}}$) which can be very efficiently solvated by water molecules. Thus the variation of the side-chain can affect structural preferences significantly with respect to solvation.

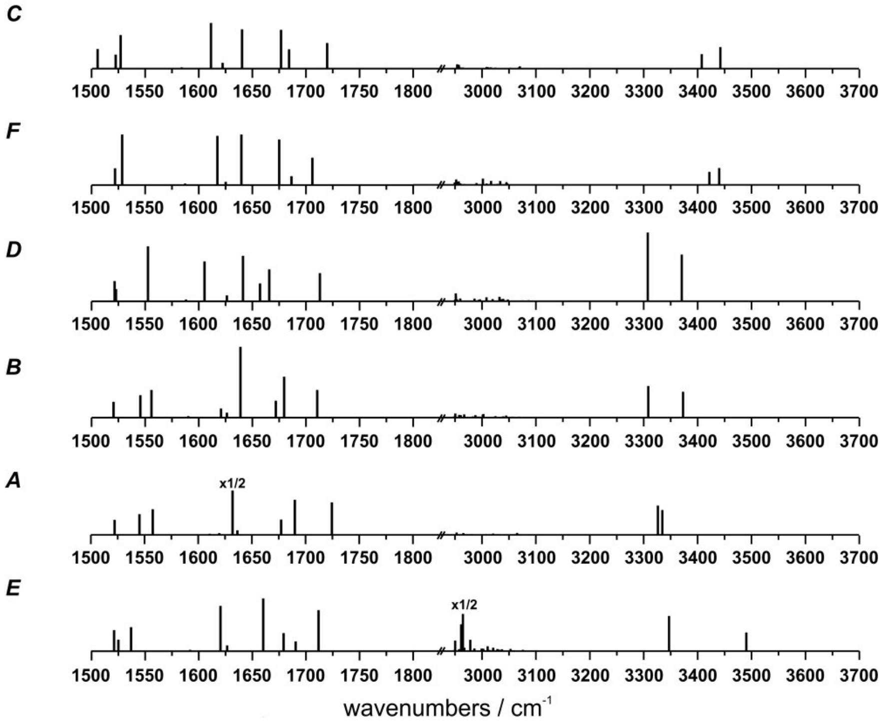


Figure S21. Calculated frequencies obtained from DFT calculations (B3LYP-D3/TZVP, ε = 78.39) of selected conformers of $\mathbf{CP_{Glu}}$ (*cf.* Figure S20). Mode-specific scaling factors of 0.9563 for hydrogen-bonded NH stretching vibrations, 0.9608 for non-hydrogen-bonded NH stretching vibrations and 0.985 for the amide I-III region were used. For a better illustration the calculated intensities in the amide A region were doubled.

f) NMR characterization of **CP**$_{\mathbf{Glu}}$

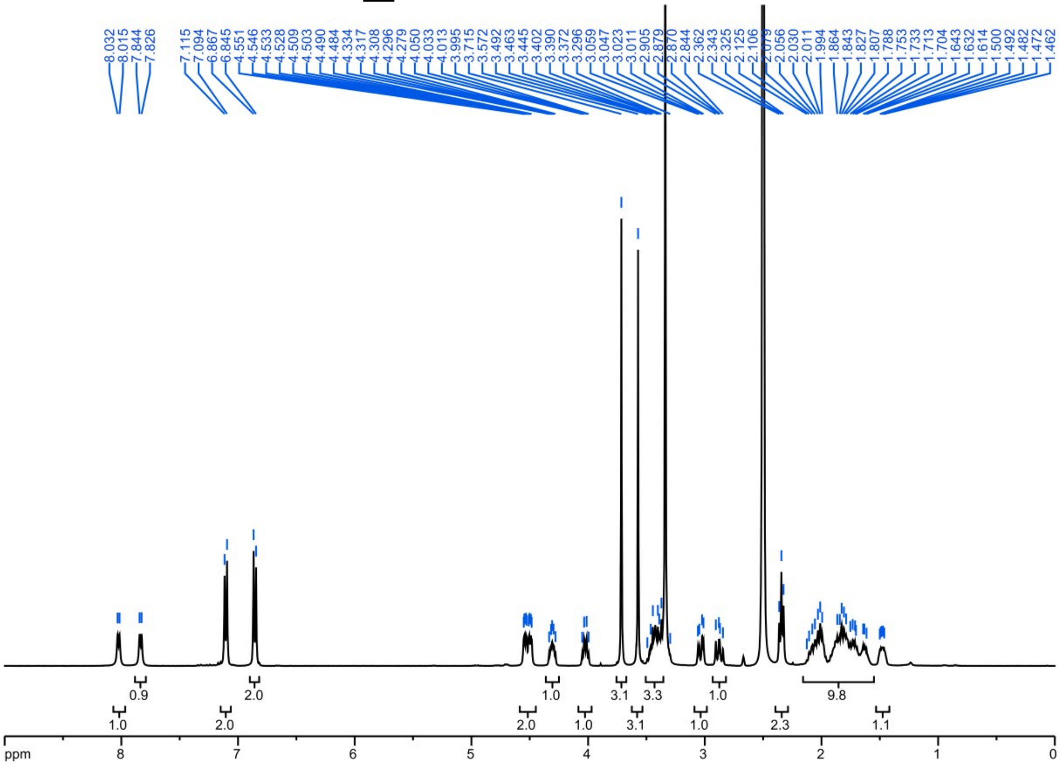


Figure S22: ^{1}H-NMR (400 MHz, DMSO-d$_{6}$) of **CP**$_{\mathbf{Glu}}$.

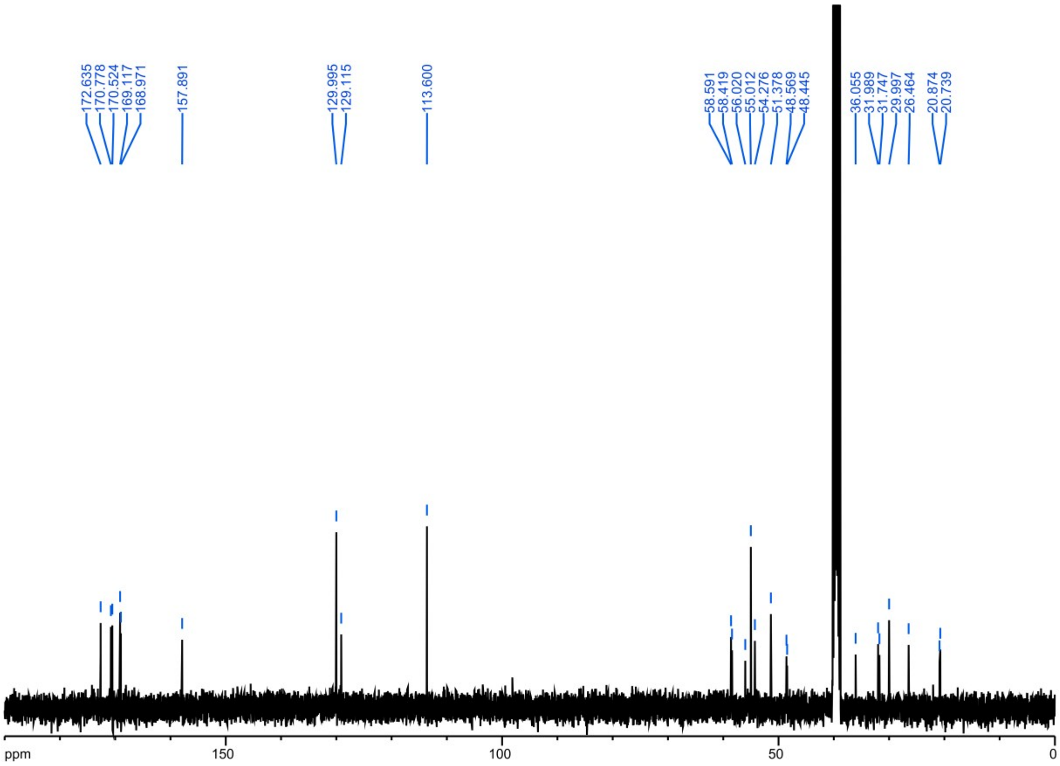


Figure S23: ^{13}C-NMR (101 MHz, DMSO-d$_{6}$) of **CP**$_{\mathbf{Glu}}$.

g) Mass spectrum of **CP$_{Glu}$**

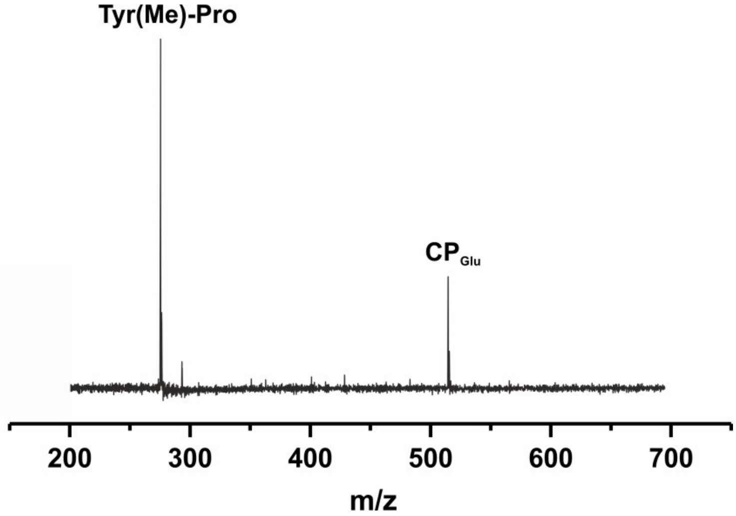


Figure S24: Mass spectrum of **CP$_{Glu}$**. The parent mass signal of **CP$_{Glu}$** obtained at a sample oven temperature of about 150 °C is constant while the IR/R2PI spectra are recorded.

References

[1] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.

[2] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *WIREs Comput Mol Sci*, 2014, **4**, 91–100.

[3] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

[4] Y. Bouteiller, J. C. Poully, C. Desfrancois, G. Gregoire, *J. Phys. Chem. A*, 2009, **113**, 6301–6307.

[5] A. Schäfer, A. Klamt, D. Sattel, J. C. W. Lohrenz, F. Eckert, *Phys. Chem. Chem. Phys.* 2000, **2**, 2187–2193.

[6] A. Klamt; G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, **5**, 799–805.

[7] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, B. M. V. Barone, 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, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision D.01, Gaussian Inc., Wallingford CT, 2013.