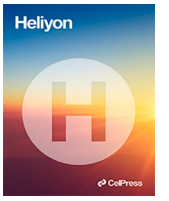
Heliyon 10 (2024) e33159

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Neighbor effect on conformational spaces of alanine residue in azapeptides

Ho-Jin Lee[j,k,*], Shi-Wei Liu[a], Máté Sulyok-Eiler[b,c], Veronika Harmat[b,d], Viktor Farkas[b,d], Zoltán Bánóczi[e,f,**], Mouna El Khabchi[g], Hua-Jun Shawn Fan[a,***], Kimihiko Hirao[h], Jong-Won Song[i]

[a] College of Chemical Engineering, Sichuan University of Science and Engineering, Zigong City, Sichuan Province, 64300, PR China
[b] Laboratory of Structural Biology and Chemistry, Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary
[c] Hevesy György PhD School of Chemistry, Eötvös Loránd University, Budapest, Hungary
[d] HUN-REN - ELTE Protein Modeling Research Group, Budapest, Hungary
[e] Department of Organic Chemistry, Institute of Chemistry, ELTE Eötvös Loránd University, 1117, Budapest, Hungary
[f] HUN-REN-ELTE Research Group of Peptide Chemistry, 1117, Budapest, Hungary
[g] LIMAS, Faculty of Sciences Dhar El Mahraz, University Sidi Mohamed Ben Abdallah, Fez, Morocco
[h] Fukui Institute for Fundamental Chemistry, Kyoto University, Takano, Nishihiraki-cho 34-4, Sakyo-ku, Kyoto, 606-8103, Japan
[i] Department of Chemistry Education, Daegu University, Daegudae-ro 201, Gyeongsan-si, Gyeongsangbuk-do, 38453, Republic of Korea
[j] Division of Natural and Mathematics Sciences, LeMoyne-Own College, Memphis, TN, 38126, USA
[k] Department of Natural Sciences, Southwest Tennessee Community College, Memphis, TN, 38015, USA

ARTICLE INFO

Keywords:
Azapeptide
Foldamer
βII-turn
βI-turn
DFT functionals
LCgau-BOP
And Lcgau-BOP+LRD

ABSTRACT

The conformational properties of Alanine (Ala) residue have been investigated to understand protein folding and develop force fields. In this work, we examined the neighbor effect on the conformational spaces of Ala residue using model azapeptides, Ac-Ala-azaGly-NHMe (**3**, **AaG**), and Ac-azaGly-Ala-NHMe (**4, aGA1**). Ramachandran energy maps were generated by scanning (φ, ψ) dihedral angles of the Ala residues in models with the fixed dihedral angles (φ = ±90°, ψ = ±0° or ±180°) of azaGly residue using LCgau-BOP and LCgau-BOP + LRD functionals in the gas and water phases. The integral-equation-formalism polarizable continuum model (IEF-PCM) and a solvation model density (SMD) were employed to mimic the solvation effect. The most favorable conformation of Ala residue in azapeptide models is found as the polyproline II (β_P), inverse γ-turn (γ'), β-sheet (β_S), right-handed helix (α_R), or left-handed helix (α_L) depending on the conformation of neighbor azaGly residue in isolated form. Solvation methods exhibit that the Ala residue favors the β_P, δ_R, and α_R conformations regardless of its position in azapeptides **3** and **4** in water. Azapeptide **5**, Ac-azaGly-Ala-NH_2 (**aGA2**), was synthesized to evaluate the theoretical results. The X-ray structure showed that azaGly residue adopts the polyproline II (β_P) and Ala residue adopts the right-handed helical (α_R) structure in **aGA2**. The conformational preferences of **aGA2** and the dimer structure of **aGA2** based on the X-ray structure were examined to assess the performance of DFT functionals. In addition, the local minima of azapeptide **6**, Ac-Phe-azaGly-NH_2 (**FaG**), were compared with the previous experimental results. SMD/LCgau-BOP + LRD

* Corresponding author. Division of Natural and Mathematics Sciences, LeMoyne-Own College, Memphis, TN, 38126, USA.
** Corresponding author. Department of Organic Chemistry, Institute of Chemistry, ELTE Eötvös Loránd University, 1117, Budapest, Hungary.
*** Corresponding author. College of Chemical Engineering, Sichuan University of Science and Engineering, Zigong City, Sichuan Province, 64300, PR China.
E-mail addresses: hjinlab@hotmail.com (H.-J. Lee), zoltan.banoczi@ttk.elte.hu (Z. Bánóczi), fan27713@yahoo.com (H.-J. Shawn Fan).

https://doi.org/10.1016/j.heliyon.2024.e33159
Received 25 April 2024; Received in revised form 10 June 2024; Accepted 14 June 2024
Available online 15 June 2024

methods agreed well with the reported experimental results. The results suggest the importance of weak dispersion interactions, neighbor effect, and solvent influence in the conformational preferences of Ala residue in model azapeptides.

1. Introduction

Conformational properties of Alanine residue in peptides have been investigated extensively with theories [1–20] and experiments [21–24] to understand the protein folding and development of the force fields [7,25–27]. Theoretical studies reported that Ala dipeptide models, Ac-Ala-NH_2 and Ac-Ala-NHMe, favor the inverse γ-turn (γ') and extended (ε or β_S) conformation in the gas phase (Fig. 1a and c) but prefer polyproline II(β_P) conformer in water [1–20]. Density-functional theory(DFT) calculations predicted the high level of the β_P population of Ac-Ala-NHMe surrounded by six molecules of water, which implies the impact of water on the conformational preferences of Ala residue [22]. In line with theoretical studies, several experimental results supported that Ala residue in dipeptide models, such as Ac-Ala-NH_2 [21] and Ac-Ala-NHMe (**1**) (Fig. 1a) [24], adopt a mixture of γ' and the β_S in the gas phase (Fig. 1c). Infrared (IR) and Raman studies of the conformational population of blocked Ala dipeptide with acetyl and N-methyl group have shown a high β_P population (60 %), followed by extended β_S conformer (29 %), and right-handed helix (α_R) conformation (11 %) in water [28]. Parchansky et al. [29] recorded Raman optical activity(ROA) spectra of Ac-Ala-NHMe, showing that Ala dipeptide adopts the mixture of α_R and β_P conformers in water. The conformational behaviors of Ala residues in tripeptide models were also investigated [30–33]. Chahkandi et al. [32] examined the conformations of For-Ser-Ala-NH_2 using B3LYP/6-311 + G(d,p) and M06-2X/6-311 + G(d,p), reporting that Ala residue favors γ' and γ conformers in the isolated form. Kalvoda et al. [31] reported detailed mapping of the conformational space of tripeptide models, such as Ac-Xaa-Yaa-NHMe, where Xaa and Yaa are any amino acid residues. They performed single-point DFT calculations at the BP86-D3/Dgauss-DZVP//COSMO-RS level in water based on the optimized molecular geometries with the GFN2-xTB semiempirical method. They showed that the Ala residue in Ala-Gly favors the β_S, δ_R, and α_R conformers, and that of the Gly-Ala sequence favors β_S, β_P, and α_R conformers depending on the conformation of Gly residue. Prasad et al. reported a diverse data set, PEPCONF, of peptide conformational energies calculated by the LC-ωPBE-XDM/aug-cc-pVTZ method for single-point energy calculations [33]. The results implied that Ala residue in peptide might adopt specific conformations. It

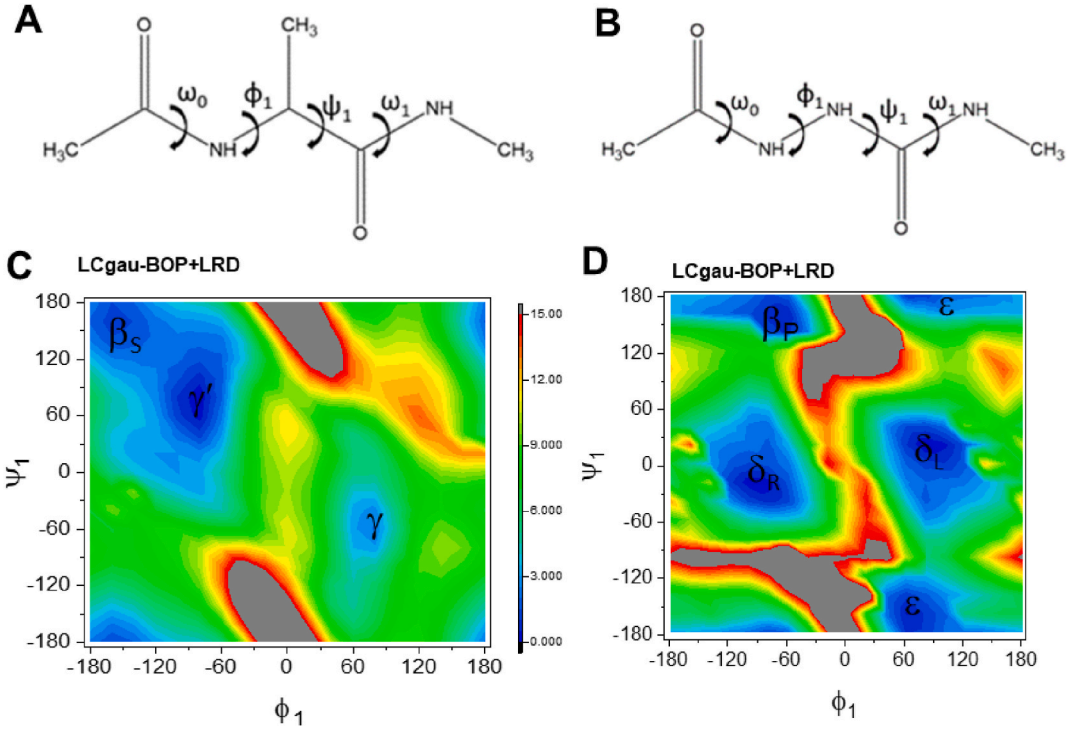


Fig. 1. Chemical structure and dihedral angle nomenclature of Ac-Ala-NHMe (**1**) and Ac-azaGly-NHMe (**2**) peptide models, respectively (**a**, **b**). Potential energy maps of (**c**) Ac-Ala-NHMe (**1**) and (**d**) Ac-azaGly-NHMe (**2**), respectively, at the LCgau-BOP + LRD/6-31 + G(d) level in the gas phase. Geometry optimization of all variables except (φ_1, ψ_1) was performed on a grid with 15° spacing. The dihedral angle of ω_0 and ω_1 was set to ~180°. The energy area from the lowest energy (blue) to 15 kcal/mol (red) is shown. The Ala residue favors $\beta_S(C_5)$ and γ' regions; the azaGly residue favors the δ_R and δ_L, β_P, or ε regions. The nomenclature of conformers was adopted from Karplus et al. [34] (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

has been reported that Ala residue in the coil library adopts the five conformers, such as γ', β_S, β_P, α_R, and α_L [25]. Many lines of evidence showed that Ala residue favors the specific conformers or ensemble of the minima depending on the environment and nearest neighbor residue. However, the neighbor and solvent effects on the conformational preference of Ala residue in the peptide are still limited due to the flexibility of peptides.

Peptidomimetics mimic natural peptides, which modify the interesting backbone or side chain [35–40]. The modification in the peptide backbone restricts peptide conformations, reducing the degrees of freedom [41]. Thus, we reasoned that peptidomimetics would be an excellent instrument for understanding the intrinsic backbone preferences of amino acid residue in peptides. In this work, we chose azaglycine (azaGly), which is the replacement of the $C_\alpha H$ group of glycine residue by an N atom (Fig. 1b). Azaglycine residue can adopt the mainly four conformers δ_R, ($\varphi = -90° \pm 30°$, $\psi = 0° \pm 30°$), δ_L ($\varphi = +90° \pm 30°$, $\psi = 0° \pm 30°$), β_P ($\varphi = -90° \pm 30°$, $\psi = 180° \pm 30°$), and ε ($\varphi = +90° \pm 30°$, $\psi = 180° \pm 30°$) (Fig. 1d) [41–46]. Previously, researchers rationalized that incorporating azaGly residue would stabilize the β-turn structure [47,48]. Indeed, the NMR(nuclear magnetic resonance) structure of Ac-Phe-azaGly-NH_2 adopted β-II(primary) and β-I turn(minor) structure in DMSO solvent [47]. Yan et al. [49] also reported that the X-ray and NMR structure of the series of azapeptides containing thioamide moiety, showing that Ac-Ala-azaGly-ψ[CSNH]-Ph and Ac-Phe-azaGly-ψ[CSNH]-Ph adopted the β-II turn mostly. The results suggest that the i+1 Ala residue prefers the polyproline II (β_P) to different conformers when azaGly adopts the δ_L conformer in the polar solvent. However, it remains unknown how Ala residue's conformational preferences would affect neighbor residue conformations or be affected by the near neighbor residue and solvent.

To investigate the neighbor effect on the conformational properties of Ala residue in peptide, we employed two azapeptides, Ac-Ala-azaGly-NHMe (**3**, **AaG**) (Fig. 2a) and Ac-azaGly-Ala-NHMe (**4**, **aGA1**) (Fig. 2b), using DFT functionals including the LCgau-BOP [50], and LCgau-BOP + LRD [51,52], the LCgau-BOP combined with the original local response dispersion (LRD) [53]. In the previous reports, LCgau-BOP + LRD reproduced inter- and intramolecular dispersion energies with reduced errors, which is neglected in the conventional DFT functionals [54,55]. We generated the potential energy Ramachandran plots of Ala residues in model peptides with the fixed dihedral angle of azaGly to (φ,$\psi = \pm 90°$, 0° or 180°) at the LCgau-BOP and LCgau-BOP + LRD level of theories in the gas phase and water. Polarizable continuum models (IEF-PCM and SMD) were used to describe the water. Based on the potential energy surface of Ala residues in model peptides, we could identify the local minima of azapeptides **3** and **4** in the gas phase and water. For comparison, we used hybrid B3LYP functional and B3LYP with Grimme's D3 dispersion effect (B3LYP-D3) [56] to calculate the local minima of azapeptides **3** and **4**. The most favorable conformers of azapeptides **3** and **4** in the gas phase and water are summarized (SI Tables S1-S6). Furthermore, we could synthesize azapeptide **5**, Ac-azaGly-Ala-NH_2 (**aGA2**), and obtain its X-ray structure in the solid phase and circular dichroism(CD) spectrum in water. Using time-dependent density-functional theory (TD-DFT) calculation, we could predict the electronic circular dichroism (ECD) spectrum for the most favorable conformers of azapeptide **5**. To evaluate our theoretical results, we examined the conformational behaviors of azapeptide **6**, Ac-Phe-azaGly-NH_2 (**FaG**), compared with the previous experimental results. Our results will provide knowledge of the intrinsic conformational properties of Ala residue in peptides and guide a new foldamer using azaglycine.

2. Calculation and experimental methods

Gaussian 09/16 programs [57,58] were used to investigate the conformational behaviors of azapeptides. Density functional theory (DFT) functionals, such as LCgau-BOP and LCgau-BOP + LRD, were used. In brief, the LCgau-BOP functional is that the LC and LCgau schemes were applied to the Becke 1988 exchange [59] + one-parameter progressive (OP) [60] correlation functional. We used the three fitted parameters ($\mu = 0.42$, $a = 0.011$, and $k = 18.0$) in LCgau-BOP calculations [61]. The LCgau-BOP functional was combined with the original local response dispersion(LRD) method with newly optimized parameters [62]. In addition, we used B3LYP and B3LYP-D3 functionals in comparison. Ramachandran energy maps of Ala dipeptide, Ac-Ala-NHMe (**1**), and azaGly peptide,

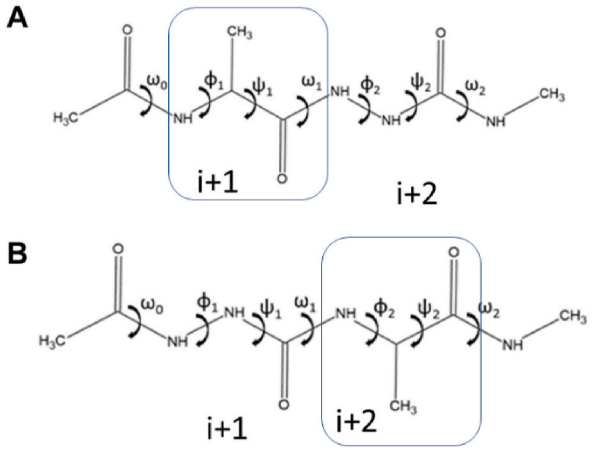


Fig. 2. Chemical structures and dihedral angle nomenclature of (**a**) Ac-Ala-azaGly-NHMe (**3, AaG**) and (**b**) Ac-azaGly-Ala-NHMe (**4, aGA1**).

Ac-azaGly-NHMe (**2**), were generated by scanning of (φ,ψ) torsional angle of Ala or azaGly residue ranging from −180° to +180° in steps of 15° (625 conformers) at the LCgau-BOP + LRD/6-31 + G(d) level of theory (Fig. 1). To examine the conformational preferences of Ala residues in tripeptides, Ac-Ala-azaGly-NHMe (**3, AaG**) (Fig. 2a) and Ac-azaGly-Ala-NHMe (**4, aGA1**) (Fig. 2b), we also generated Ramachandran plots at the LCgau-BOP-LRD/6-31 + G(d) level of theory. During the scanning of (φ,ψ) dihedral angle of Ala residue ranging from 180° to +180° in steps of 15° (625 conformers), the (φ, ψ) dihedral angles of azaGly were fixed to (φ = ±90°, ψ = 0° or 180°) in the gas phase and water. To estimate the effect of water on the topology of the energy surfaces, we partly optimized the model tripeptides in water using IEF-PCM(integral equation formalism variant of the Polarizable Continuum Model) [63] and SMD [64] solvation models at the LCgau-BOP and LCgau-BOP + LRD functional.

We selected the starting structures using two methods to characterize the local minima for azapeptides **3** and **4**. The starting structures of azapeptides were selected (1) based on the Ramachandran energy maps of Ala residues (Figs. 3–6) and (2) the conformational search using HyperChem 8.0 [65] at the RM1 semiempirical method [66]. These initial structures were optimized fully at LCgau-BOP, LCgau-BOP + LRD, B3LYP, and B3LYP-D3 with a 6–311++G(2d,2p) basis set in the gas phase and water. Furthermore, we synthesized azapeptide **5**, Ac-azaGly-Ala-NH_2 (**aGA2**), and computed the most favored conformers of azapeptide **5** at the same level of theories, compared with the X-ray structure of azapeptide **5** (SI Tables S13-S15). The electronic circular dichroism (ECD) spectrum for the conformers of azapeptide **5** was computed using time-dependent-density-functional theory (TD-DFT) to compare the experiment circular dichroism(CD) spectrum of azapeptide **5** (SI Tables S11-S13 and Figures S2 and S3). To evaluate the theoretical results, we also calculated the minima conformations of Ac-Phe-azaGly-NH_2 (**6**) and compared them with the previously reported experimental results (SI Tables S14-S18). Frequencies were computed at the same level of theories to confirm each stationary point as a minimum. Conformation populations of model azapeptides **3**–**6** were calculated at 298K based on the relative energies [24]. The figures were

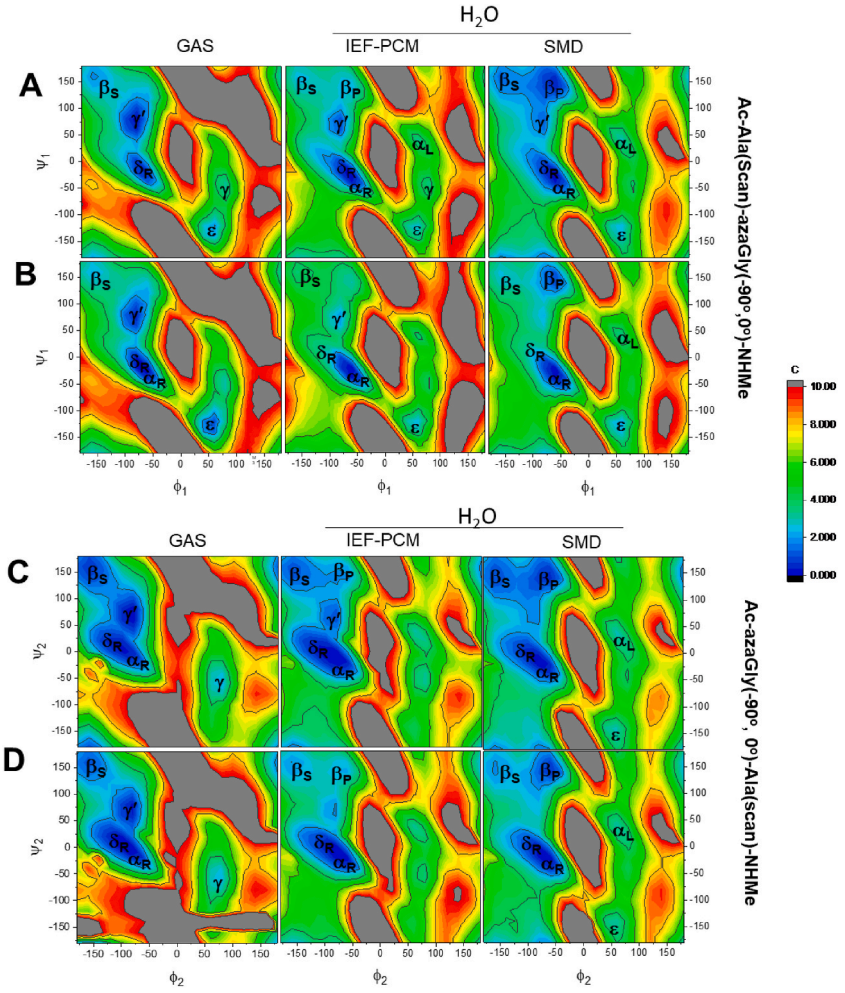


Fig. 3. Ramachandran energy maps of Ac-Ala(φ_1,ψ_1 = scan)-azaGly ($\varphi_2 = -90°$, $\psi_2 = 0°$)-NHMe (**A** and **B**) and Ac-azaGly($\varphi_1 = -90°$, $\psi_1 = 0°$)-Ala (φ_2,ψ_2 = scan)-NHMe (**C**, **D**) at the LCgau-BOP (**A** and **C**) and LCgau-BOP + LRD (**B** and **D**) with the 6-31 + G(d) basis set. Geometry optimization of all variables except (φ_1, ψ_1) for **AaG** or (φ_2, ψ_2) for **aGA1** was performed on a grid with 15° spacing. The dihedral angle of ω_0 and ω_2 was ~180°. The energy area from the lowest energy (blue) to 10 kcal/mol (red) is shown. IEF-PCM and SMD solvation methods were used to consider the water environment. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

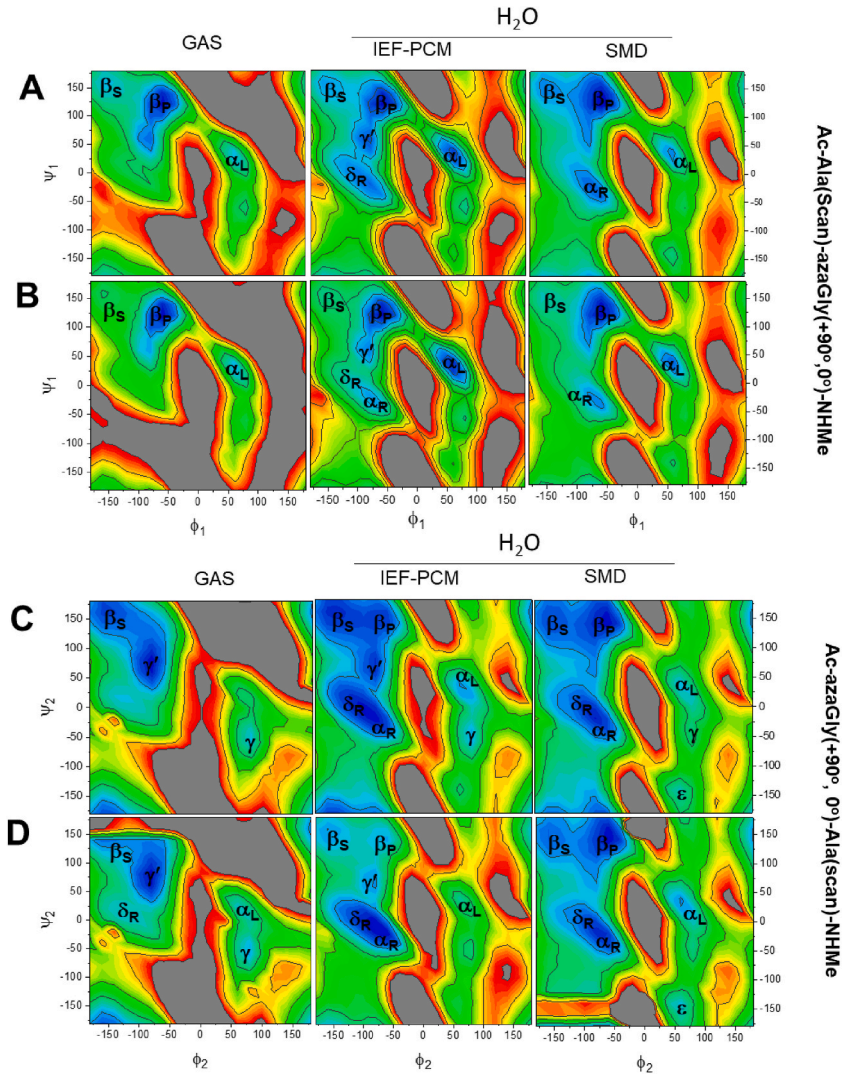


Fig. 4. Ramachandran energy maps of Ac-Ala(φ_1,ψ_1 = scan)-azaGly($\varphi_2 = +90°$, $\psi_2 = 0°$)-NHMe (**A** and **B**) and Ac-azaGly($\varphi_1 = +90°$, $\psi_1 = 0°$)-Ala (φ,ψ = scan)-NHMe (**C** and **D**) at the LCgau-BOP (**A** and **C**) and LCgau-BOP + LRD (**B** and **D**) with the 6-31 + G(d) basis set. Geometry optimization of all variables except (φ_1, ψ_1) for **AaG** or (φ_2, ψ_2) for **aGA** was performed on a grid with 15° spacing. The dihedral angle of ω_0 and ω_2 was ~180°. The energy area from the lowest energy (blue) to 10 kcal/mol (red) is shown. IEF-PCM and SMD methods were used to consider the water environment. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

generated using GaussView 6 [67] and Origin programs. Multiwfn [68] and VMD programs generated the non-covalent interactions (NCI) plot for the **aGA2-05** and **aGA2-12** conformers.

Synthesis of Ac-azaGly-Ala-NH_2 (5, aGA2). Azapeptide **5** was synthesized on Rink Amid MBHA-resin by solid phase peptide synthesis using the Fmoc/*t*Bu strategy. The Fmoc-Ala-OH was coupled to resins using diisopropyl carbodiimide (DIC) and Oxyma pure (ethyl cyano(hydroxyimino)acetate) coupling reagents. The azaglycine was built as previously described in Ref. [69]. The acetyl group was introduced using acetic anhydride. The acetylated dipeptide was cleaved from the resin by TFA using deionized water and triisopropylsilane as scavengers. The crude peptide was purified by RP-HPLC and was characterized by analytical RP-HPLC and ESI MS.

Crystallographic study of Ac-azaGly-Ala-NH_2 (5, aGA2). Transparent needle shape crystals of azapeptide 6 were grown from water solution. X-ray diffraction data from a single crystal (dimensions: 0.06 × 0.06 × 0.77 mm) were collected at 105.3 K on a Rigaku XtaLab Synergy-R diffractometer using Cu-K_α radiation (λ = 1.54184 Å). Data reduction was done using the diffractometer software [CrysAlisPro 1.171.40.14e (Rigaku OD, 2018)]. Olex2 v1.2 [70] was used for molecular graphics program and structure determination. Azapeptide **5** crystallized in the monoclinic space group $P2_1$, with one molecule within the asymmetric unit (Fig. 9). The phase problem was solved using the intrinsic method with the SHELXT program [71]. The structure was refined by full-matrix least-squares techniques (SHELXL-2014/7) on F^2 [72] Hydrogen atoms were refined in the riding positions (with bond lengths and angles constrained to ideal geometry), except for H4, H5, and H8 (see Fig. 9a for atom numbering): Position of difference electron density peaks of H atoms bound to nitrogen atoms of the azaGly and Ala moieties indicated that configuration around N4, N5, and N8 may

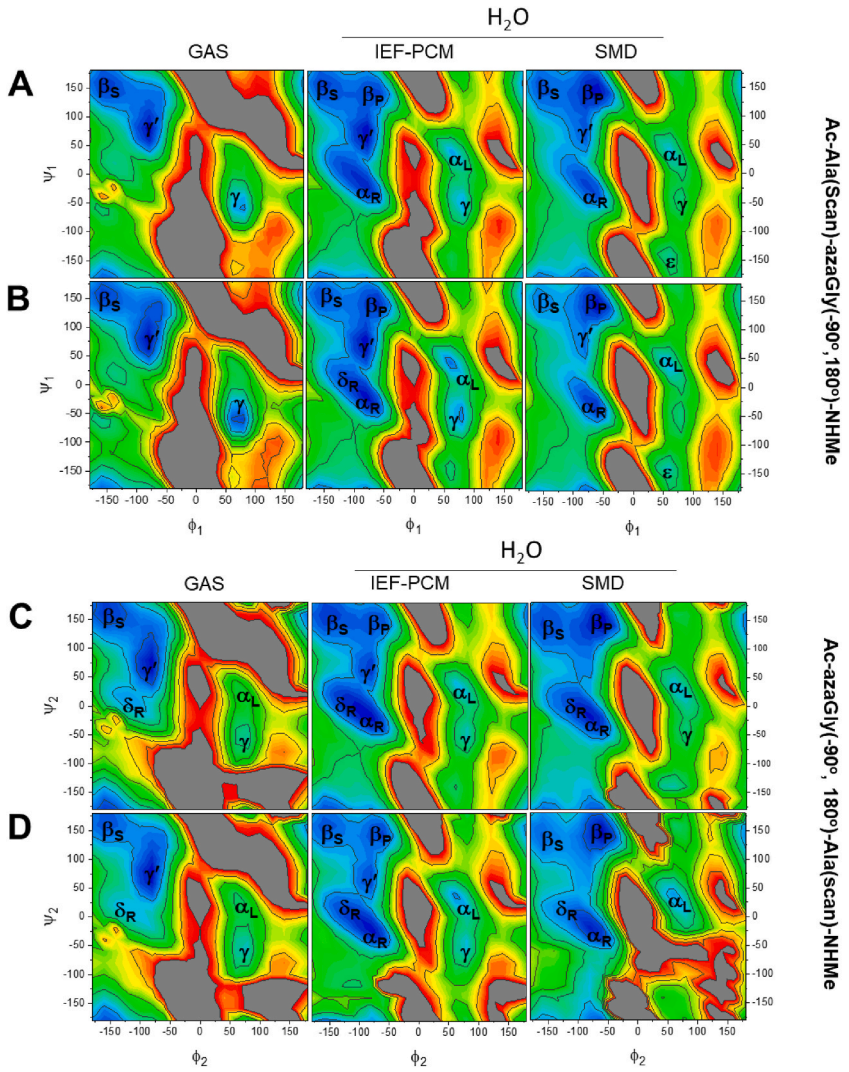


Fig. 5. Ramachandran energy maps of Ac-Ala(φ_1,ψ_1 = scan)-azaGly($\varphi_2 = -90°$, $\psi_2 = 180°$)-NHMe (**A** and **B**) and Ac-azaGly($\varphi_1 = +90°$, $\psi_1 = 180°$)-Ala(φ_2,ψ_2 = scan)-NHMe (**C** and **D**) at the LCgau-BOP (**A** and **C**) and LCgau-BOP + LRD (**B** and **D**) with the 6-31 + G(d) basis set. Geometry optimization of all variables except (φ_1, ψ_1 = scan) for **AaG** or (φ_2, ψ_2 = scan) for **aGA1** was performed on a grid with 15° spacing. The dihedral angle of ω_0 and ω_2 was ~180°. The energy area from the lowest energy (blue) to 10 kcal/mol (red) is shown. IEF-PCM and SMD solvation methods were used to consider the water effect. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

significantly deviate from planarity so H4, H5, and H8 were assigned to these peaks and were refined as free atoms. The crystallographic parameters, data collection, and structure refinement details are summarized in Tables S9 and S10. The structures were analyzed, and the figures were created using the Mercury program [73]. Validation was carried out using CheckCIF/PLATON [74]. The Crystal structure of azapeptide **5** was deposited with the Cambridge Crystallographic Data Centre and can be obtained free of charge with CCDC deposition number **CCDC 2290581.**

Electronic circular dichroism spectroscopy. Far-UV ECD spectrum was recorded on a Jasco J1500 spectrophotometer using a cuvette with a path length of 1.0 mm with a peptide concentration of 2.5 mmol in water. Typical spectral accumulation parameters were a scan rate of 50 nm min^{-1} with a 1 nm bandwidth and a 0.2 nm step resolution over the 185–260 nm wavelength range with five scans averaged at 5 °C. The temperature in the cell was controlled by a Peltier-type heating system. The solvent reference spectra were used as baselines automatically subtracted from the peptide spectra. The raw ellipticity data were converted into mean residue molar ellipticity units ($[\Theta]_{MR}$, deg × cm^2 × $dmol^{-1}$).

3. Results and discussion

We generated Ramachandran energy maps of Ala residue at i+1 or i+2 position in model tripeptides, Ac-Ala-azaGly-NHMe (**3**) and

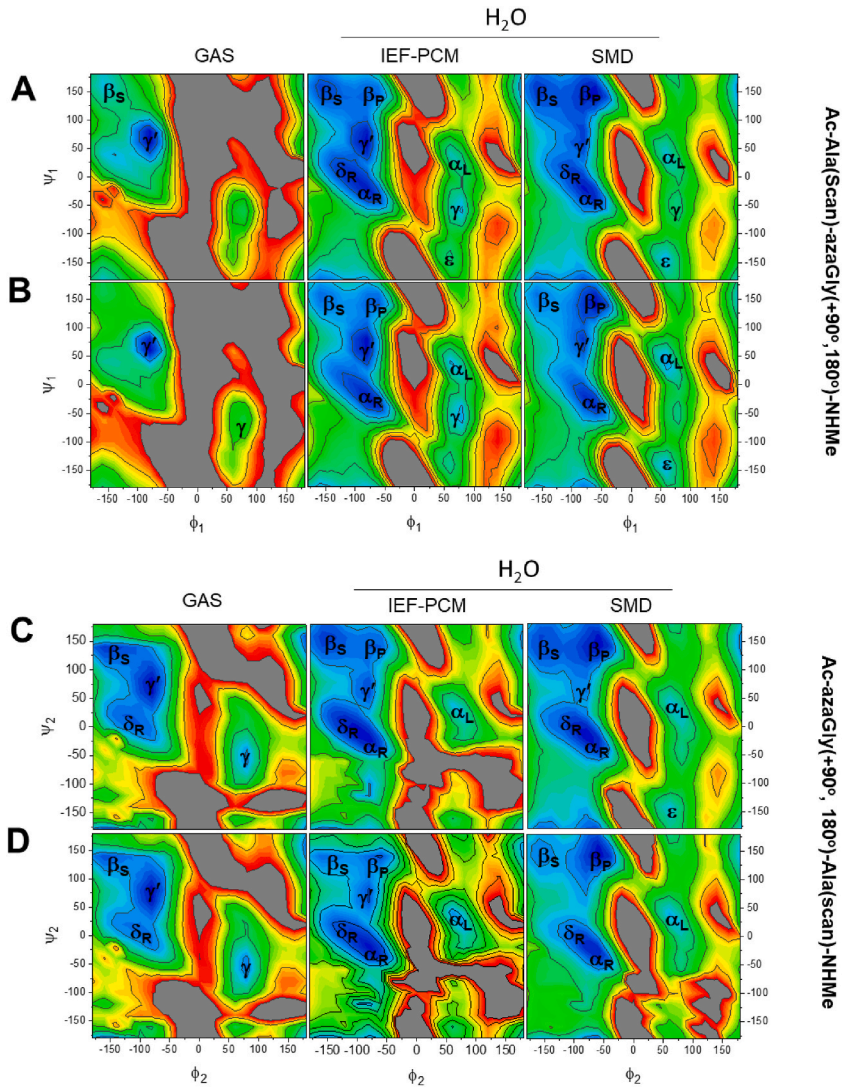


Fig. 6. Ramachandran energy maps of Ac-Ala(φ_1,ψ_1 = scan)-azaGly ($\varphi_2 = +90°$, $\psi_2 = 180°$)-NHMe (**A** and **B**) and Ac-azaGly($\varphi_1 = +90°$, $\psi_1 = 180°$)-Ala(φ_2,ψ_2 = scan)-NHMe (C, D) at the LCgau-BOP (**A** and **C**) and LCgau-BOP + LRD (**B** and **D**) with the 6-31 + G(d) basis set. Geometry optimization of all variables except (φ_1, ψ_1 = scan) for **AaG** or (φ_2, ψ_2 = scan) for **aGA1** was performed on a grid with 15° spacing. The dihedral angle of ω_0 and ω_2 was set to ~180°. The energy area from the lowest energy (blue) to 10 kcal/mol (red) is shown. IEF-PCM and SMD solvation methods were used to consider the water effect. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Ac-azaGly-Ala-NHMe (**4**), as the dihedral angle of azaGly residue was fixed, such as right-handed and left-handed bridged ($\varphi = -90°$, $\psi = 0°$ or $\varphi = +90°$, $\psi = 0°$), polyproline II ($\varphi = -90°$, $\psi = 180°$) and extended ($\varphi = +90°$, $\psi = 180°$) conformations at the LCgau-BOP and LCgau-BOP + LRD with the 6-31 + G(d) basis set in the gas and water phase (Figs. 3–6). We employed IEF-PCM and SMD solvation methods to consider the water effect. Tables 1 and 2 contain the structural information of local minima for azapeptides **3** and **4** (SI Tables S1-S6) at the LCgau-BOP, LCgau-BOP + LRD, B3LYP, and B3LYP-D3 functions with the 6–311++G(2d,2p) basis set in the gas and water phase.

3.1. Ramachandran energy maps of Ala residue when azaGly adopted δ_R conformation ($\varphi = -90°$, $\psi = 0°$)

Fig. 3a and b shows the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy maps of i+1 Ala residue of azapeptide **3** as the i+2 azaGly residue has a biased conformer (**$\varphi_2 = -90°$, $\psi_2 = 0°$**) in the gas phase (left) and water (center and right). In the isolated form, LCgau-BOP and LCgau-BOP + LRD functionals predict that the inverse gamma-turn (γ') and β-strand (β_S) and the bridged (δ_R) regions of the i+1 Ala residue are stable (Fig. 3a, left). Because the α_R and δ_R regions were not found as local minima for Ala dipeptide (Fig. 1c) [19,26], the δ_R and α_R regions could be affected by the biased conformation of neighbor azaGly residue in the gas

Table 1
The backbone dihedral angles and relative energies (ΔE in kcal/mol) of the most favorable conformer of Ac-Ala-azaGly-NHMe (**3**, **AaG**) calculated at the LCgau-BOP, LCgau-BOP-LRD, B3LYP, and B3LYP-D3 functionals with 6–311++G(2d,2p) basis set in isolated form and in water.

LCgau-BOP									
No.	conformer	ω_0	φ_1	ψ_1	ω_1	φ_2	ψ_2	ω_2	ΔE
AaG02	$\beta_P\delta_L$	172.3	−59.6	129.4	−176.8	70.0	21.1	−177.5	0.00
AaG05	$\gamma'\varepsilon$	169.0	−79.9	77.0	−159.7	63.4	−163.8	176.0	0.04
AaG03	$\gamma'\delta_R$	−176.7	−83.4	80.0	−167.6	−95.2	−22.1	176.9	0.31
LCgau-BOP + LRD									
AaG05	$\gamma'\varepsilon$	168.0	−79.4	78.0	−157.5	60.7	−162.6	175.9	0.00
AaG02	$\beta_P\delta_L$	170.2	−59.2	130.0	−175.5	66.1	23.0	−176.3	0.01
AaG01	$\delta_R\delta_R$	−167.9	−70.4	−18.8	174.4	−69.4	−22.0	176.5	1.44
B3LYP									
AaG03	$\gamma'\delta_R$	−177.6	−82.3	80.2	−167.8	−99.5	−21.3	178.1	0.00
AaG02	$\beta_P\delta_L$	175.0	−62.8	124.5	−175.7	73.9	20.4	−178.9	0.19
AaG05	$\gamma'\varepsilon$	172.0	−80.8	75.0	−163.6	67.6	−164.2	177.2	0.63
B3LYP-D3									
AaG02	$\beta_P\delta_L$	173.1	−63.9	121.9	−173.2	68.4	23.3	−178.1	0.00
AaG05	$\gamma'\varepsilon$	166.8	−79.3	77.4	−157.3	62.6	163.5	176.6	0.44
AaG01	$\delta_R\delta_R$	−169.8	−69.7	−17.7	172.7	−70.7	−21.5	177.4	1.13
SMD/LCgau-BOP									
AaG02	$\beta_P\delta_L$	178.8	−56.7	133.8	−179.6	72.4	17.9	−177.8	0.00
AaG01	$\alpha_R\delta_R$	−174.0	−63.7	−25.3	178.8	−73.9	−19.8	179.5	0.31
AaG10	$\beta_P\delta_L$	178.6	−55.8	134.6	−176.1	98.2	−17.0	−174.1	0.50
SMD/**LCgau-BOP + LRD**									
AaG02	$\beta_P\delta_L$	175.5	−55.5	135.3	−177.7	70.2	19.0	−176.2	0.00
AaG01	$\alpha_R\delta_R$	−175.5	−61.9	−28.1	177.8	−70.8	−21.2	177.2	0.31
AaG10	$\beta_P\delta_L$	177.4	−55.7	134.2	−175.8	93.9	−14.6	−172.1	0.88
SMD/B3LYP									
AaG02	$\beta_P\delta_L$	179.6	−62.2	134.8	176.4	77.8	18.1	−178.5	0.00
AaG01	$\alpha_R\delta_R$	−175.8	−65.5	−24.2	178.0	−76.5	−18.7	179.7	0.31
AaG10	$\beta_P\delta_L$	−179.8	−58.6	130.0	−176.0	99.8	−16.9	−174.2	0.31
SMD/B3LYP-D3									
AaG01	$\alpha_R\delta_R$	−178.5	−62.1	−26.8	176.1	−72.4	−19.9	178.6	0.00
AaG02	$\beta_P\delta_L$	178.9	−59.9	129.8	−176.1	68.6	19.2	−176.7	0.19
AaG10	$\beta_P\delta_L$	179.0	−59.2	129.5	−174.2	95.5	−15.8	−174.0	0.82

phase. The α_R region of i+1 Ala residue could be stabilized by the intramolecular hydrogen bond with the i+2 azaGly residue that adopted the δ_R conformer, which forms the β-I turn structure. LCgau-BOP and LCgau-POP + LRD functional show that region β_S of i+1 Ala residue has a higher energy surface than δ_R region in the isolated form (Fig. 3a and b, left).

IEF-PCM/LCgau-BOP method predicts that γ' and α_R/δ_R regions of the i+1 Ala residue are stable (Fig. 3a, center). Interestingly, including the dispersion correction (+LRD) predicts that the γ' and β_S regions have a higher energy (Fig. 3b, middle). The prediction of SMD/LCgau-BOP shows the β_S, β_P, γ' and δ_R/α_R regions found to be local minima in water (Fig. 3a, right). Interestingly, the inclusion of the dispersion correction (+LRD) predicts that the polyproline II(β_P) and δ_R regions are only stabilized in water (Fig. 3b, right). The results indicate that the δ_R region of the i+1 Ala residue in azapeptide **3** is stabilized by the biased conformation of the i+2 azaGly residue, which is related to the neighbor residue. The preference of β_P region of i+1 Ala residue in azapeptide **3** would be related to the solvation effect (Fig. 3b, right).

Fig. 3c and d are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy map of the i+2 position of Ala residue in azapeptide **4** as the i+1 azaGly adopts the δ_R conformer ($\varphi_1 = -90°$, $\psi_1 = 0°$) in the gas and water phase. LCgau-BOP predicts that the β_S and γ' regions, likely Ala dipeptide, and the α_R and δ_R regions of the i+2 Ala residue are stable in isolate form (Fig. 3c, left). The δ_L and α_L regions of the i+2 Ala residue are on a higher energy surface than the α_R and δ_R regions of the i+2 Ala residue of azapeptide **4**. Remarkably, the favored α_R and δ_R regions of the i+2 position Ala residue would result from the intramolecular hydrogen bond with the biased δ_R conformation of the i+1 azaGly residue. The LCgau-BOP + LRD Ramachandran energy map of the i+2 Ala residue (Fig. 3d, left) is like the LCgau-BOP method's results (Fig. 3c, left). This indicates that the effect of dispersion correction is minor in this case.

The IEF-PCM/LCgau-BOP method also shows similar conformational preferences of the i+2 Ala residue in the gas phase, except the β_P region in water (Fig. 3c, middle). However, the IEF-PCM/LCgau-BOP + LRD method predicts that the α_R and δ_R regions, but not the β_p and β_S regions, are found to be stable in water (Fig. 3d, middle). SMD/LCgau-BOP results are slightly different from IEF-PCM/LCgau-BOP results. SMD/LCgau-BOP predicts that the preference of i+2 Ala residue is found in the β_P, β_S, α_R, and δ_R region, but not γ'region, in water (Fig. 3c, right). The results of the SMD/LCgau-BOP + LRD method are like those of IEF-PCM/LCgau-BOP + LRD, but SMD/LCgau-BOP + LRD predicts α_R region as the most stable, followed by the δ_R and β_P regions of the i+2 position Ala in water (Fig. 3d, right). In sum, the β_P region of Ala residue may be stabilized by water regardless of the position of Ala residue. The α_R (or δ_R)

Table 2
The backbone dihedral angles and relative energies (ΔE in kcal/mol) of the most favorable conformer of Ac-azaGly-Ala-NHMe (**4**, **aGA**) calculated at the LCgau-BOP, LCgau-BOP-LRD, B3LYP, and B3LYP-D3 functionals with 6–311++G(2d,2p) basis set in isolated form and in water.

No	conformer	ω_0	φ_1	ψ_1	ω_1	φ_2	ψ_2	ω_2	ΔE
LCgau-BOP									
aGA11	$\delta_R\gamma'$	−169.5	−93.6	−21.3	−179.0	−82.8	75.5	−172.8	0.00
aGA03	$\delta_L\gamma'$	168.9	91.6	20.3	−179.0	−85.2	73.5	−174.6	0.10
aGA05	$\alpha_R\delta_R$	−170.0	−71.4	−22.6	177.6	−93.0	6.5	175.1	0.41
LCgau-BOP + LRD									
aGA11	$\delta_R\gamma'$	−175.3	−74.4	−30.5	−164.2	−80.0	83.9	−167.1	0.00
aGA05	$\alpha_R\delta_R$	−171.4	−70.4	−23.7	176.4	−88.1	3.3	175.4	0.06
aGA03	$\delta_L\gamma'$	169.8	85.3	20.9	−179.0	−85.1	74.2	−173.8	0.69
B3LYP									
aGA11	$\delta_R\gamma'$	−169.7	−96.4	−20.7	−178.6	−82.5	75.9	−173.6	0.00
aGA03	$\delta_L\gamma'$	169.6	93.8	19.7	179.0	−83.9	75.1	−174.8	0.12
aGA05	$\alpha_R\delta_R$	−170.7	−71.0	−22.3	176.9	−94.4	6.6	175.3	0.65
B3LYP-D3									
aGA05	$\alpha_R\delta_R$	−172.8	−69.4	−23.6	174.6	−88.6	3.7	175.7	0.00
aGA11	$\delta_R\gamma'$	−176.0	−74.0	−30.9	163.5	−80.6	81.8	−165.7	0.16
aGA03	$\delta_L\gamma'$	171.2	86.9	20.1	178.8	−83.8	75.2	−173.7	0.82
SMD/**LCgau-BOP**									
aGA05	$\alpha_R\alpha_R$	−174.8	−70.2	−21.7	173.0	−67.1	−19.7	178.9	0.00
aGA06	$\delta_R\beta_S$	−176.5	−76.4	−21.5	−179.7	−154.5	163.7	176.0	1.26
aGA04	$\varepsilon\beta_S$	−176.4	67.4	−159.6	−170.9	−160.5	147.1	174.8	1.48
SMD/LCgau-BOP + LRD									
aGA05	$\alpha_R\alpha_R$	−174.8	−70.2	−21.7	173.0	−67.1	−19.7	178.9	0.00
aGA11	$\delta_R\beta_P$	−179.5	−67.8	−32.0	159.3	−80.6	114.7	−177.4	1.95
aGA06	$\delta_R\beta_S$	−176.5	−76.4	−21.5	−179.7	−154.5	163.7	176.0	2.40
SMD/B3LYP									
aGA05	$\delta_R\delta_R$	−175.4	−72.3	−20.8	174.7	−75.8	−14.2	180.0	0.00
aGA03	$\delta_L\beta_P$	173.9	83.1	18.7	175.1	−65.4	142.5	174.6	0.25
aGA01	$\beta_P\beta_P$	−179.0	−77.4	162.6	162.6	−63.2	143.5	175.4	0.38
SMD/B3LYP-D3									
aGA05	$\delta_R\delta_R$	−176.7	−69.3	−22.5	172.6	−72.1	−15.3	179.3	0.00
aGA06	$\delta_R\beta_S$	−178.1	−75.4	−20.8	−176.5	−156.6	153.1	174.5	2.70
aGA11	$\delta_R\gamma'$	−177.9	−71.0	−25.8	166.4	−79.6	86.5	−170.0	2.71

conformation of the Ala residue would be related to the biased δ_R conformation of the neighbor azaGly residue.

3.2. *Ramachandran energy maps of Ala residue when azaGly adopted δ_L conformation ($\varphi = +90°$, $\psi = 0°$)*

Fig. 4a and b are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy map of the i+1 Ala residue in azapeptide **3** as the i+2 position of azaGly reside adopted the δ_L ($\varphi = +90°$, $\psi = 0°$) in the gas phase and water. In the gas phase, both methods predict that the β_P region of the i+1 position of Ala residue is the most stable (Fig. 4a, left), which was not found in Ala dipeptide (Fig. 1c). The result indicated that the β_P region of the i+1 Ala residue is stabilized through the intramolecular hydrogen bond with the i+2 azaGly residue. The β_S, γ', and α_L regions have a higher energy surface than the β_P region of the i+1 Ala residue in the gas phase. Interestingly, Including the long-range dependency (+LRD) stabilizes the β_P region but destabilizes the β_S and γ' regions of i+1 Ala residue of azapeptide **3** in the gas phase (Fig. 4b, left). In water, the β_S, γ', β_P, δ_R, and α_R are found to be local minima (Fig. 4a, center) at the IEF-PCM/LCgau-BOP function. Including LRD predicts that β_P regions are the most stable, followed by α_L regions of i+1 Ala residue (Fig. 4b center). Note that SMD/LCgau-BOP prediction is different from IEF-PCM/LCgau-BOP prediction. The γ' region is destabilized, and the β_S and α_R regions of i+1 Ala residue are located at the local minima at the SMD/LCgau-BOP (Fig. 4a, right). Including the long-range dependency (+LRD) would also predict the β_P region of the i+1 Ala residue but not the β_S region in water (Fig. 4b, right). All calculations show that the neighbor effect stabilizes the β_P region of the i+1 Ala, and the solvent stabilizes the other stable areas, such as the δ_R and α_L regions.

Fig. 4c and d are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy maps of the i+2 position Ala residue in azapeptide **4** as the i+1 azaGly is set to the conformer δ_L ($\varphi = +90°$, $\psi = 0°$) in the gas and water phase. The Ramachandran energy plot of the i+2 position Ala residue is like that of the Ala dipeptide in Fig. 1c, showing that the regions β_S and γ' are expected to be most stable (Fig. 4c, left). The result suggests that the biased i+1 azaGly residue would not affect the conformational preferences of the i+2 Ala residue in azapeptide **4**. In addition, including LRD does not change the conformational behaviors of the i+2 Ala residue in the gas phase (Fig. 4d, left). In water, the β_P, γ', δ_R, and α_R regions of the i+2 Ala residue are found in the local energy minima at the IEF-PCM/LCgau-BOP prediction. The results indicate that water might stabilize the β_P, δ_R, and α_R regions but not the β_S region of the i+2 Ala residue (Fig. 4c, center). SMD/LCgau-BOP calculation predicts that β_P, δ_R, and α_R regions, but not γ' region of the i+2 Ala residue, are

energetically stable. The results of SMD/LCgau-BOP + LRD Ramachandran energy map of the i+2 Ala residue (Fig. 4c, right) is like that of SMD/LCgau-BOP calculation, indicating that the long-range correction might be negligible in the conformational preference of the i+2 Ala residue in water. In sum, the β_P region of the i+1 Ala residue would be stabilized by the biased i+2 azaGly residue. The regions β_P, $\alpha_{R,}$ and δ_R of the i+2 Ala residue would be stabilized by solvent, not neighbor residue.

3.3. Ramachandran energy maps of Ala residue when azaGly adopted β_P conformation ($\varphi = -90^\circ$, $\psi = 180^\circ$)

Fig. 5a and b are the respective LCgau-BOP and LCgau-BOP + LRD Ramachan energy map of the i+1 position Ala residue as the i+2 position azaGly adopts the β_P conformer in the gas phase and water. In the isolated form, i+1 Ala residue favors the γ' and β_S regions energetically, but γ regions have higher energy surfaces than the γ' region (Fig. 5a, left). The inclusion of LRD shows similar conformational preferences of i+1 Ala residue, except for γ regions (Fig. 5b, left), indicating that the dispersion correction may be negligible. In the water phase, both solvation methods predict that the β_P and δ_R regions are the local minima (Fig. 5a, center and right). SMD/LCgau-BOP calculation indicates that the β_P and α_R regions of i+1 Ala residue are most stable in energy, but γ' their areas are destabilized (Fig. 5b, right). In addition, SMD/LCgau-BOP + LRD calculation indicates that the i+1 Ala favors the β_P region but not the γ' region in the water phase. The results suggest that the solvation effect would be the preference of (β_P region of the i+1 Ala residue with the biased i+2 azaGly residue.

Fig. 5c and d are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy map of i+2 Ala residue for azapeptide **4** as the i+1 azaGly adopts the β_P conformer in the gas phase and water. In the gas phase, the LCgau-BOP method predicts that the i+2 Ala residue favors γ' and β_S regions, which are the most stable (Fig. 5c, left). Interestingly, LCgau-BOP calculation shows that the area δ_R of the i+2 Ala residue is in a higher energy surface. In the water phase, the IEF-PCM and SMD solvent methods with the LCgau-BOP method predict that the regions β_P, δ_R, and α_R of the i+2 Ala residue are stable in water (Fig. 5c, middle and right). Note that the solvent environment would stabilize the α_R region of i+2 Ala residue (Fig. 5c, middle; and Fig. 5d, middle). The SMD/LCgau-BOP method indicates that the γ' region of the i+2 Ala residue is destabilized in the water phase. IEF-PCM(SMD)/LCgau-BOP + LRD results are like IEF-PCM(SMD)/LCgau-BOP (Fig. 5c and d, right). The results imply that the effect of LRD on conformational preferences for the i+2 Ala residue in azapeptide **4** is negligible. Both solvation models with LCgau-BOP + LRD functional predict that β_S, β_P, and α_R regions are most stable in water (Fig. 5d, middle and right). Note that the SMD/LCgau-BOP + LRD method predicts that the α_R region, but not the δ_R region, is stable (Fig. 5c, right). In sum, the neighbor effect of the β_P conformer in azaGly on Ala residue may be minor, and water may stabilize the β_P and α_R regions of Ala residues regardless of the position. The results imply that the effect of LRD on conformational preferences for i+2 position Ala residue in azapeptide **4** is negligible.

3.4. Ramachandran energy maps of Ala residue when azaGly adopted ε conformation ($\varphi = +90^\circ$, $\psi = 180^\circ$)

Fig. 6a and b are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy map of i+1 Ala residue as the i+2 azaGly adopts the ε conformation ($\varphi = +90^\circ$, $\psi = 180^\circ$) in azapeptide **3** in the gas phase and water. Unlike the Ala dipeptide, the γ' region of i+1 Ala residue is the most stable in isolated form (Fig. 6a, left). The β_S and γ areas are also observed as local minima with higher energy than the γ' region. The results suggest that conformational behaviors of the i+1 Ala residue for azapeptide **3** are affected by the neighbor's biased conformation ($\varphi = +90^\circ$, $\psi = 180^\circ$) of the i+2 azaGly residue. Including LRD does not change the overall energy surface of i+1 Ala residue for azapeptide **3** in the gas phase (Fig. 6b, left), compared to LCgau-BOP (Fig. 6a left) results.

IEF-PCM/LCgau-BOP calculation shows that the regionsγ', β_S, β_P, δ_R, and α_R of i+1 Ala residue are stabilized (Fig. 6a, middle). Including LRD stabilizes the regions γ' and α_R of the i+1 Ala residue in water (Fig. 6b, middle). SMD/LCgau-BOP method predicts that the β_P region is the most stable, but γ' region is destabilized (Fig. 6a, right). Note that the ε conformer of the i+2 azaGly affects the conformational preference of the i+1 position of Ala residue. The solvent model predicts that the β_P and α_R conformers of the i+1 Ala residue might be stabilized by water.

Fig. 6c and d are the respective LCgau-BOP and LCgau-BOP + LRD Ramachandran energy map of the i+2 position Ala residue for azapeptide **4** as i+1 azaGly residue adopts the ε conformation ($\varphi = +90^\circ$, $\psi = 180^\circ$) at the i+1 position in the gas phase and water. In the gas phase, LCgau-BOP calculation shows that the regions β_S and γ' of the i+2 Ala residue are the most stable, remarkably, and δ_R region is also found as local minima. The results show that the conformational preferences of the i+2 position A are affected by the biased conformer of the i+1 azaGly residue. The inclusion of LRD does not change the conformational behaviors of the i+2 position of Ala residue in the gas phase (Fig. 6d, left). In water, both methods show similar potential energy maps of the i+2 Ala residue (Fig. 6c middle and right). The IEF-PCM/LCgau-BOP calculation indicates that the regions δ_R and α_R of the i+2 Ala residue are the most stable (Fig. 6c, middle). SMD solvent model predicts that the β_P and α_R regions are the most stable in water. Including LRD predicts that the α_R region of the i+2 Ala residue is stabilized for both solvent models (Fig. 6d, middle and right). IEF-PCM/LCgau-BOP + LRD calculation predicts that the areas β_S, β_P, and γ' The local minima of the i+2 Ala residue are in a higher energy surface (Fig. 6d, middle). SMD/LCgau-BOP + LRD shows the stabilization of the β_P region but not the β_S region in water (Fig. 6d, right). The results indicate that dispersion correction (+LRD) affects the conformation preferences of the i+2 Ala residue in water. In sum, the conformational behavior of the i+2 Ala residue is affected by the biased ε conformer of the i+1 position of azaGly. The water may stabilize the β_P and α_R conformer of i+2 Ala residue when azaGly adopts the extended conformation at the i+1 position.

3.5. Minima conformations of azapeptides 3 and 4

To characterize the conformational behaviors of azapeptides **3** and **4**, we examined the minima using four DFT functionals,

including LCgau-BOP, LCgau-BOP + LRD, B3LYP, and B3LYP-D3 in the gas phase and water with 6–311++G(2d,2p) basis set. Tables 1 and 2 summarize the top 3 ranked conformers for azapeptides **3** and **4** (Figs. 7a and 8a). The identified 10 or 12 minima for azapeptides **3** and **4** are listed in SI Tables S1-S6. The conformation population of each conformer is shown in Fig. 7b and 8b and listed in SI Tables S7 and S8. The following are the details of local minima for the model azapeptides **3** and **4**, respectively.

General trends of Ac-Ala-azaGly-NHMe (3, AaG): The most preferred conformer for azapeptide **3** differs depending on the methods used. In the gas phase, DFT functionals used predict **AaG02** ($\beta_P\delta_L$), **AaG03**($\gamma'\delta_R$), or **AaG05** ($\gamma'\varepsilon$) conformer as the most stable conformers (Table 1 and Fig. 7a). LCgau-BOP method predicts two conformers **AaG02** ($\beta_P\delta_L$) and **AaG05** ($\gamma'\varepsilon$) are found in the same energy, showing that 32 % of **AaG02**, 34 % of **AaG05**, and 21 % of **AaG03**. Including LRD increases the population of **AaG02** and **AaG05** by 12 % and 11 %, respectively, though populations of other conformers decreased (SI Table S7). This trend is valid for the B3LYP-D3 method, showing that the three conformers **AaG02**, **AaG03**, and **AaG05** were found in a less than 1 kcal/mol range. The dispersion correction increases the population of **AaG02** by 26 % and **AaG05** by 12 % in the isolated form (Table 1 and SI Table S7). The results suggest that the i+1 position Ala residue favors the polyproline II(β_P) and inverse-gamma turn(γ') conformers in the isolated form. Since inverse-gamma turn is also found in the Ala dipeptide model, the preferred polyproline conformer of i+1 Ala depends on the conformation of i+2 azaGly residue.

We investigated the solvent effect on the conformational preferences of azapeptide **3**. The IEF-PCM model indicates that all DFT functionals used predict **AaG01**($\alpha_R\delta_R$) as the lowest energy conformation (Table 1, SI Table S2), corresponding to the β-I turn structure. LCgau-BOP shows 46 % of **AaG01**, and including LRD increases the population of **AgG01** by 11 %. Compared to IEF-PCM/LCgau-BOP results, IEF-PCM/B3LYP shows a reduced population of **AaG01**(31 %) and **AaG02** (23 %). B3LYP-D3 functional shows the increment population of **AaG01** by 34 % in the isolated form. Generally, the prediction of the IEF-PCM model indicates that the β-I turn structure is stabilized in the water phase (SI Table S7). However, the SMD/LCgau-BOP method predicts **AaG02** (β-II) as the lowest energy conformation (Table 1 and SI Tables S3 and S7). Including LRD shows similar relative stability among conformers but increases the population of **AaG01** ($\delta_R\delta_R$) and **AaG02** ($\beta_P\delta_L$) by 6 % and 9 %, respectively. SMD/B3LYP predicts 31 % of **AaG02** and 20 % of **AaG10**, corresponding to the β-II turn structures. The inclusion of D3 empirical dispersion correction indicates **AaG01**($\alpha_R\delta_R$) conformer, corresponding to β-I-turn, as the lowest conformation. B3LYP-D3 increases the population of **AaG01** by 28 % and shows a similar population of **AaG02** and **AaG10** conformers. The results suggest that water would stabilize the α_R or β_P conformer for i+1 Ala residue.

General trends of Ac-azaGly-Ala-NHMe (4, aGA1): Table 2 lists the top-ranked three conformers of azapeptide **4** calculated at the four DFT functionals in the gas phase and water (SI Table S8). For the isolated form, LCgau-BOP predicts **aGA1-11**($\delta_R\gamma'$) as the lowest energy conformation, followed by **aGA1-03** ($\delta_L\gamma'$) and **aGA1-05** ($\alpha_R\delta_R$, β-I turn). The population of **aGA1-11** is estimated to be 31 %, and that of **aGA1-03** is 26 % (Fig. 8). The inclusion of LRD changes the relative energy stability of conformations of **aGA1**. Although

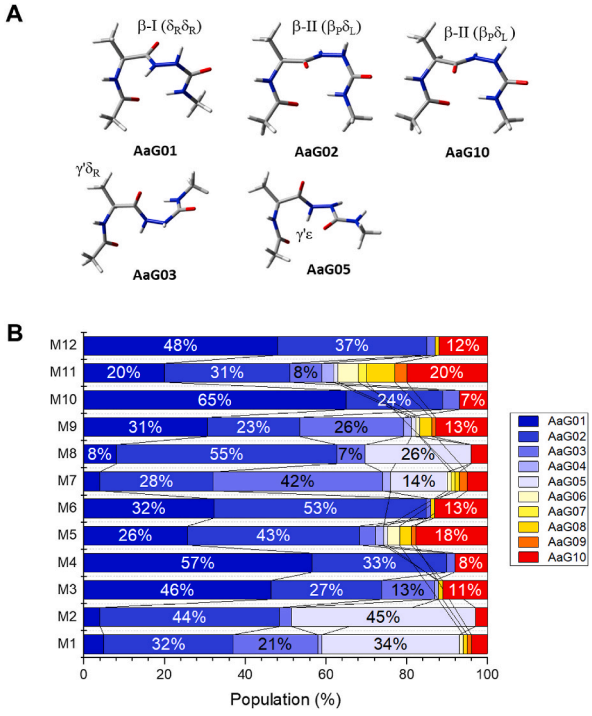


Fig. 7. The most favorable conformation of Ac-Ala-azaGly-NHMe (**3, AaG**) at the SMD/LCgau-BOP + LRD/6–311++G(2d,2p) (**A**) and population of conformers in the gas and water (**B**). For y-axis, **M1**:LCgau-BOP; **M2**:LCgau + LRD; **M3**:IEF-PCM/LCgau-BOP; **M4**:IEF-PCM/LCgau-BOP + LRD; **M5**: SMD/LCgau-BOP; **M6**: SMD/LCgau-BOP + LRD; **M7**: B3LYP; **M8**:B3LYP-D3; **M9**: IEF-PCM/B3LYP; **M10**: IEF-PCM/B3LYP-D3; **M11**: SMD/B3LYP; **M12**: SMD/B3LYP-D3 with the 6–311++G(2d,2p) basis set used.

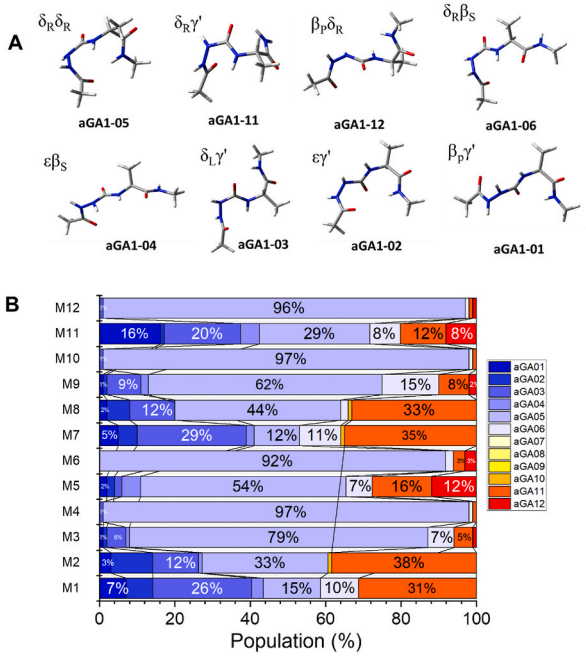


Fig. 8. The most favorable conformation of Ac-azaGly-Ala-NHMe (**4**, **aGA1**) at the SMD/LCgau-BOP + LRD/6–311++G(2d,2p) (**A**) and population of conformers for azapeptide **4** in the gas and water (**B**). For y-axis, **M1**:LCgau-BOP; **M2**:LCgau + LRD; **M3**:IEF-PCM/LCgau-BOP; **M4**:IEF-PCM/LCgau-BOP + LRD; **M5**: SMD/LCgau-BOP; **M6**: SMD/LCgau-BOP + LRD; **M7**: B3LYP; **M8**:B3LYP-D3; **M9**: IEF-PCM/B3LYP; **M10**: IEF-PCM/B3LYP-D3; **M11**: SMD/B3LYP; **M12**: SMD/B3LYP-D3 with the 6–311++G(2d,2p) basis set used.

aGA1-11($\delta_R\gamma'$) is the lowest energy conformation, corresponding to 38 % of the total population, **aGA1-05** (33 %) is the second lowest, and **aGA1-03** (12 %). The results of B3LYP are like those of LCgau-BOP. The results show that 35 % of **aGA1-11**, 29 % of **aGA1-03**, and 12 % of **aGA1-05** are estimated in the isolated form (Table 2). The inclusion of D3 empirical dispersion correction changes the relative energy order, showing that **aGA1-05** (44 %) is the lowest energy minimum, followed by **aGA1-11** (33 %) and then **aGA1-03** (12 %) in the isolated form (Table 2 and Fig. 8). The results showed that the relative populations of **aGA1** minima for LCgau-BOP + LRD are like those for B3LYP-D3 in the gas phase.

In water, the IEF-PCM method with all DFT calculations used here predicts that **aGA1-05** ($\delta_R\delta_R$, β-I turn) is the lowest energy conformation. Remarkably, the population of **aGA1-05** is 80 % at the IEF-PCM/LCgau-BOP and 96 % at the IEF-PCM/LCgau-BOP + LRD level of theories. For the IEF-PCM/B3LYP method, 62 % of **aGA1-05** is estimated. Considering the D3 method, 96 % of **aGA1-05** conformer is predicted in the water phase. The SMD/LCgau-BOP method estimates 55 % of **aGA1-05** in water. SMD/LCgau-BOP + LRD is predicted to be 90 % of **aGA1-05** in the water phase. For SMD/B3LYP, 29 % of **aGA1-05**, 20 % of **aGA1-03**, and 16 % of **aGA1-01** are estimated. However, SMD/B3LYP-D3 predicts **aGA1-05** is the lowest energy conformation, corresponding to 95 % of the total population.

In sum, **aGA1-11**($\delta_R\gamma'$) is the lowest energy conformation for most DFT functionals in the gas phase; **aGA1-05** ($\delta_R\delta_R$) is the most favorable in water. The results imply that i+2 Ala residue favors inverse γ-turn(γ′) in the gas phase and helical/bridged conformer (α_R/δ_R) in the water. The results showed that the relative populations of **aGA1** minima for SMD (or IEF-PCM)/LCgau-BOP + LRD are like those for SMD (IEF-PCM)/B3LYP-D3 in the water phase.

3.6. *X-ray crystallographic and DFT studies on Ac-azaGly-Ala-NH$_2$ (5, aGA2)*

To examine the conformational properties of i+2 Ala residue, we synthesized azapeptide **5**, Ac-azaGly-Ala-NH$_2$ (Fig. 9a). The crystal structure of azapeptide **5** shows that i+1 azaGly residue adopts polyproline II (β_P: $\varphi_1 = -66.4°$, $\psi_1 = 161.1°$) and i+2 Ala residue adopts helical or bridged conformation (δ_R: $\varphi_2 = -72.2°$, $\psi_2 = -32.9°$). Puckering of N4, N5 (azaglycine), and N8 (alanine) atoms, as well as the deviation of the acetamidyl and ureidyl moieties from planarity (SI Table S10), indicate a decrease of conjugation within these groups. Crystal packing shows a 3D network of hydrogen-bonded chains of molecules with all the donor and acceptor groups involved in intermolecular H-bonds (Fig. 9b–e, Table 3). A hydrogen-bonded chain of molecules is formed by the acetamidyl moiety of one molecule, and the ureidyl moiety of a symmetry equivalent molecule runs parallel to the diagonal of ***ab*** plane of the unit cell - the ureidyl moiety of the first molecule is involved in the same type of hydrogen-bonded chain of molecules parallel to the other diagonal (Fig. 9b). These form 2-D layers of hydrogen-bonded molecules parallel to the ***ab*** planes of the unit cell (Fig. 9c and d). The C-terminal amides form a hydrogen-bonded amide ladder along the unit cell axis ***b*** (Fig. 9d).

Note that the X-ray structure of **aGA2** in the solid differs from the conformational behaviors of **aGA1** predicted in the gas phase and

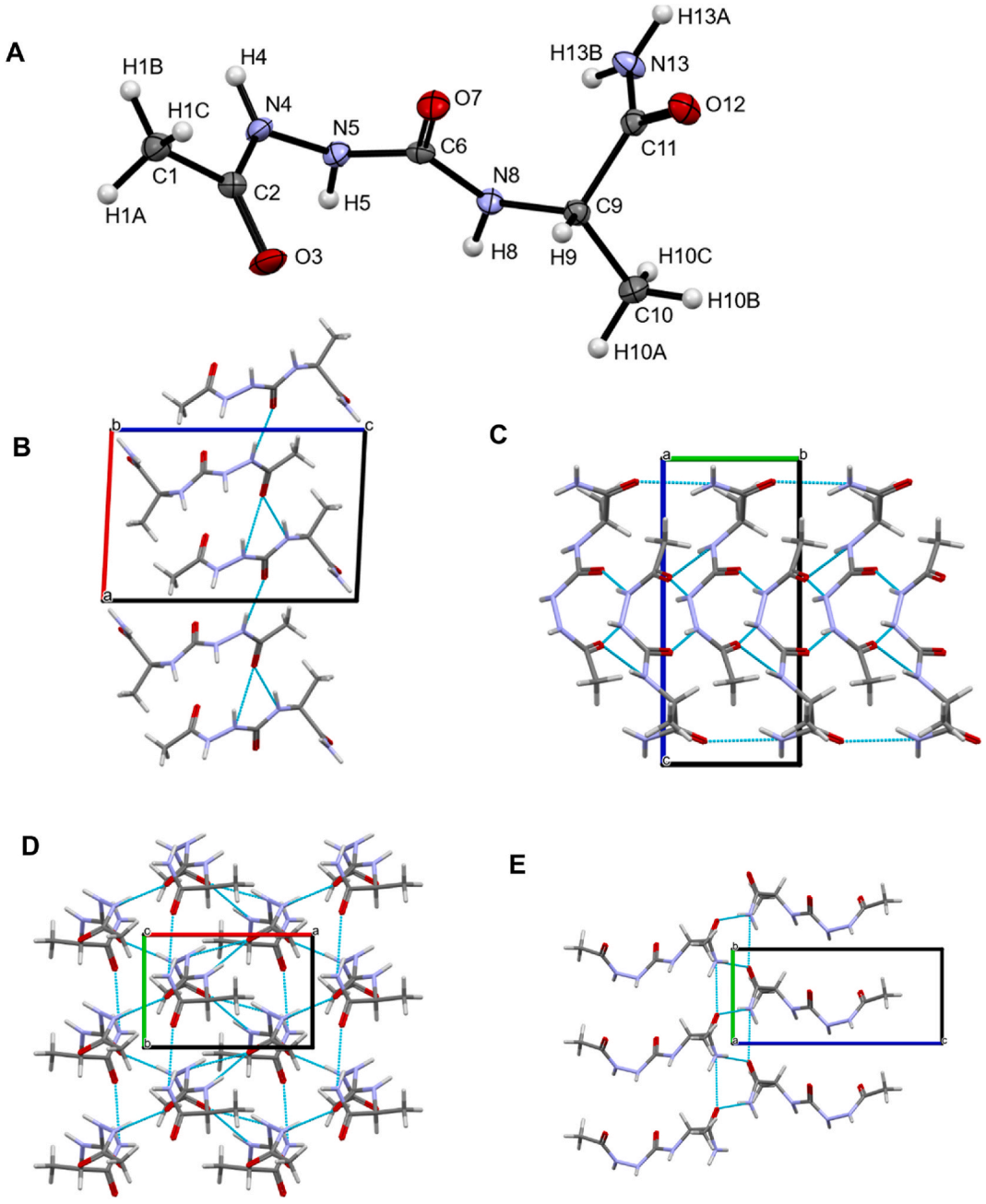


Fig. 9. (**A**) Crystal structure of Ac-azaGly-Ala-NH_2 (azapeptide **5, aGA2**). Atomic displacement ellipsoids are contoured at the 0.5 probability level. Crystal packing with H-bonds is shown as cyan dotted lines: (**B**) H-bonded chains of molecules along the diagonal of the ***ab*** plane of the unit cell (view along the ***b*** axis). (**C** and **D**) H-bonded 2-D network of molecules in the ***ab*** plane(view along unit cell axis ***a*** and ***c***, respectively). (**E**) H-bonded ladder of molecules along unit cell axis ***b*** (view along axis ***a***). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

water. The predicted **aGA1-12** ($\beta_P\delta_R$) conformer for Ac-azaGly-Ala-NHMe (**aGA1**) is located at the higher energy ($\Delta E = 2.99$–4.32 kcal/mol) for all DFT functionals in the gas phase (Table S4). In water, **aGA1-12** is found in the range of energy ($\Delta E = 1.95$–4.07 kcal/mol) at the IEF-PCM solvation model, depending on the DFT functionals. The IEF-PCM/B3LYP-D3 method predicts that **aGA1-12** is at higher energy ($\Delta E = 4.07$ kcal/mol). SMD/B3LYP-D3 solvation model seems to stabilize the **aGA1-12** conformer but is found in the range of energy ($\Delta E = 0.75$–2.64 kcal/mol) depending on the DFT calculation. Since the conformer **aGA1-12** is found at the higher energy for all DFT functionals used here, we wondered (1) whether the C-terminal methyl group may affect the conformational preferences of i+2 Ala residue (2) in the intermolecular hydrogen bond between i+1 azaGly residue may have a stabilize the specific

Table 3
Geometric parameters for hydrogen bonds formed within the crystal structure of Ac-azaGly-Ala-NH_2 (**5**, **aGA2**).

D-H.A atoms	*d*(D-H) (Å)	*d*(H.A) (Å)	<DHA (°)	*d*(D.A) (Å)	Symmetry operator for atom A
N13–H13A … O12	0.880[a]	2.101	168.52	2.969	[-*x*, *y*-1/2, -*z*][b]
N13–H13B ….O12	0.880[a]	2.160	138.51	2.879	[*x*, *y*-1, *z*][b]
C1–H1B···O7	0.980[a]	2.440	144.54	3.287	[-*x*, *y*-1/2, -*z*+1]
N5–H5···O3	0.846	2.240	146.35	2.982	[-*x*+1, *y*-1/2, -*z*+1][c]
N8–H8···O3	0.908	1.913	158.60	2.778	[-*x*+1, *y*-1/2, -*z*+1][c]
N4–H4···O7	0.852	1.943	165.69	2.777	[-*x*, *y*-1/2, -*z*+1][c]

[a] These H atoms were idealized and included as rigid groups.
[b] Forming a hydrogen-bonded ladder of molecules along unit cell axis ***d***. (Fig. 9e).
[c] Forming chains of hydrogen bonded molecules along the diagonal of unit cell plane ***ab*** (Fig. 9b–d).

conformation.

To answer these two questions, we calculated the local minima of Ac-azaGly-Ala-NH_2 (**5**, **aGA2**) using LCgau-BOP + LRD and B3LYP-D3 functionals (Fig. 10a and SI Tables S11-S13). As expected, the relative energy order among the conformers for azapeptide **5** is quite similar to that for azapeptide **4**, although the relative populations among the conformers are slightly different (SI Tables S11-S13). The lowest energy conformer for azapeptide **5** is the conformer **aGA2-05** ($\delta_R\delta_R$, β-I) with 31 % and 44 % population, at the LCgau-BOP + LRD and B3LYP-D3 functions, respectively, in the isolated form (SI Table S11). Remarkably, the second lowest energy conformer **aGA2-12** ($\beta_P\delta_R$) has a higher energy (ΔE = 3.24 kcal/mol and 4.19 kcal/mol) at the LCgau-BOP + LRD and B3LYP-D3 functional, respectively). In water, the population of the conformer **aGA2-05** is increased to 86 % and 94 % at the SMD/LCgau-BOP + LRD and SMD/B3LYP-D3 functionals, respectively (SI Table S12). SMD/LCgau-BOP + LRD and SMD/B3LYP-D3 functions stabilize the **aGA2-12** conformer (ΔE = 1.67 kcal/mol and 2.44 kcal/mol, respectively). IEF-PCM/B3LYP-D3 predicts that **aGA2-12** is the third lowest energy conformer (ΔE = 3.96 kcal/mol) (SI Table S13). The results imply that the **aGA2-12**($\beta_P\delta_R$) conformer would be stabilized by intermolecular hydrogen bonds in a solid state [43,44,75,76].

Since the X-ray structure of **aGA2** adopts the polyproline II (β_P) in the i+1 azaGly residue and α-helical (α_R) structure in the i+2 Ala residue, we assessed the performance of DFT functionals used here to see whether the conformer **aGA2-12** is stabilized through the intermolecular hydrogen bond. We calculated the dimer structures of **aGA2** using different methods (SI Figure S1). Remarkably, all DFT functionals used here failed to predict the dimer structure of **aGA2** in the isolated form (SI Fig. S1). The backbone (φ, ψ) dihedral angles of i+1 azaGly residue and i+2 Ala residue in **aGA2** in the isolated form are also quite different from those in the X-ray structure. Interestingly, IEF-PCM/LCgau-BOP + LRD, SMD/LCgau-BOP + LRD, and SMD/B3LYP-D3 calculations show a similar X-ray structure of **aGA2**. The results imply that the dispersion and solvent effect would be essential to describe the intrinsic structures of Ala residue in the azapeptide model.

3.7. CD spectrum of azapeptide 5

Although the X-ray structure of azapeptide **5** adopts the β_P conformer for azaGly and the α_R conformer for Ala residue in the solid phase, all DFT functionals used here in the gas and water predict the **aGA2-05**($\delta_R\delta_R$) conformer has the lowest energy. To characterize the conformation preference of azapeptide **5** in water, we obtained the circular dichroism (CD) spectrum of azapeptide **5** at 5 °C (Fig. 10a and b). While the CD spectrum of the peptide is sensitive to the secondary structure of protein [77], the CD spectrum of the azapeptide remains unclear. Thus, we computed the electronic circular dichroism (ECD) spectrum using time-dependent density functional theory (TD-DFT) (SI Figures S2 and S3). In addition, which DFT functionals predict the ECD spectrum of azapeptide has yet to be examined. We calculated the ECD spectrum for 12 conformers of azapeptide **5** at the SMD/TD-LCgau-BOP + LRD and SMD/TD-B3LYP-D3 functionals (SI Figures S2 and S3, Figs. 10c and d and). The computed ECD spectrum of **aGA2-05** using the SMD/TD-B3LYP-D3, but not SMD/TD-LCgau-BOP-LRD, shows good agreement with the experimental one (Fig. 10c and d). Since the predicted population of **aGA2-05** is 94 % at SMD/B3LYP-D3 level, the experiment CD spectrum of azapeptide **5** may represent the major conformer **aGA2-05** in water. The results imply that the SMD/TD-B3LYP-D3 method would help examine the ECD spectrum to explore the conformational preference of azapeptide. However, SMD/B3LYP-D3 and SMD/LCgau-BOP + LRD functionals predict that **aGA2-05** is the lowest energy conformer in water.

3.8. Assessment of DFT functionals used here using azapeptide 6, Ac-Phe-azaGly-NH_2 (6, FaG)

We found that SMD/LCgau-BOP + LRD and SMD/B3LYP-D3 methods best describe the conformational preferences of Ala residue in azapeptide **5**. We further assessed the performance of DFT functionals used in this work. We benchmarked our previous study (SI Figure S4) [47], showing that the solution structure of Ac-Phe-azaGly-NH_2 (**6, FaG**) adopted β-II (primary) and β-I (minor) turns in the polar solvent (DMSO). Thus, we re-investigated the conformational behaviors of azapeptide **6** using the LCgau-BOP + LRD and B3LYP-D3 functions in the gas phase and water using the IEF-PCM and SMD model (SI Table S14-S18). In the gas phase, LCgau-BOP + LRD and B3LYP-D3 predict that **FaG02** ($\beta_P\delta_L$(a), β-II turn) is the lowest energy conformer, corresponding to 42 % and 26 % of the population, respectively (SI Table S14). In water, IEF-PCM/LCgau-BOP + LRD and SMD/LCgau-BOP + LRD predict that the conformer **FaG02** (β–II–turn, 37–38 %) is the lowest (SI Tables S15 and S16). The conformer **FaG12** (βI'-turn, 19 % population) for IEF-PCM/SMD/LCgau-BOP + LRD and the conformer **FaG14** (βI-turn, 13 % population) for SMD/LCgau-BOP + LRD, respectively, are

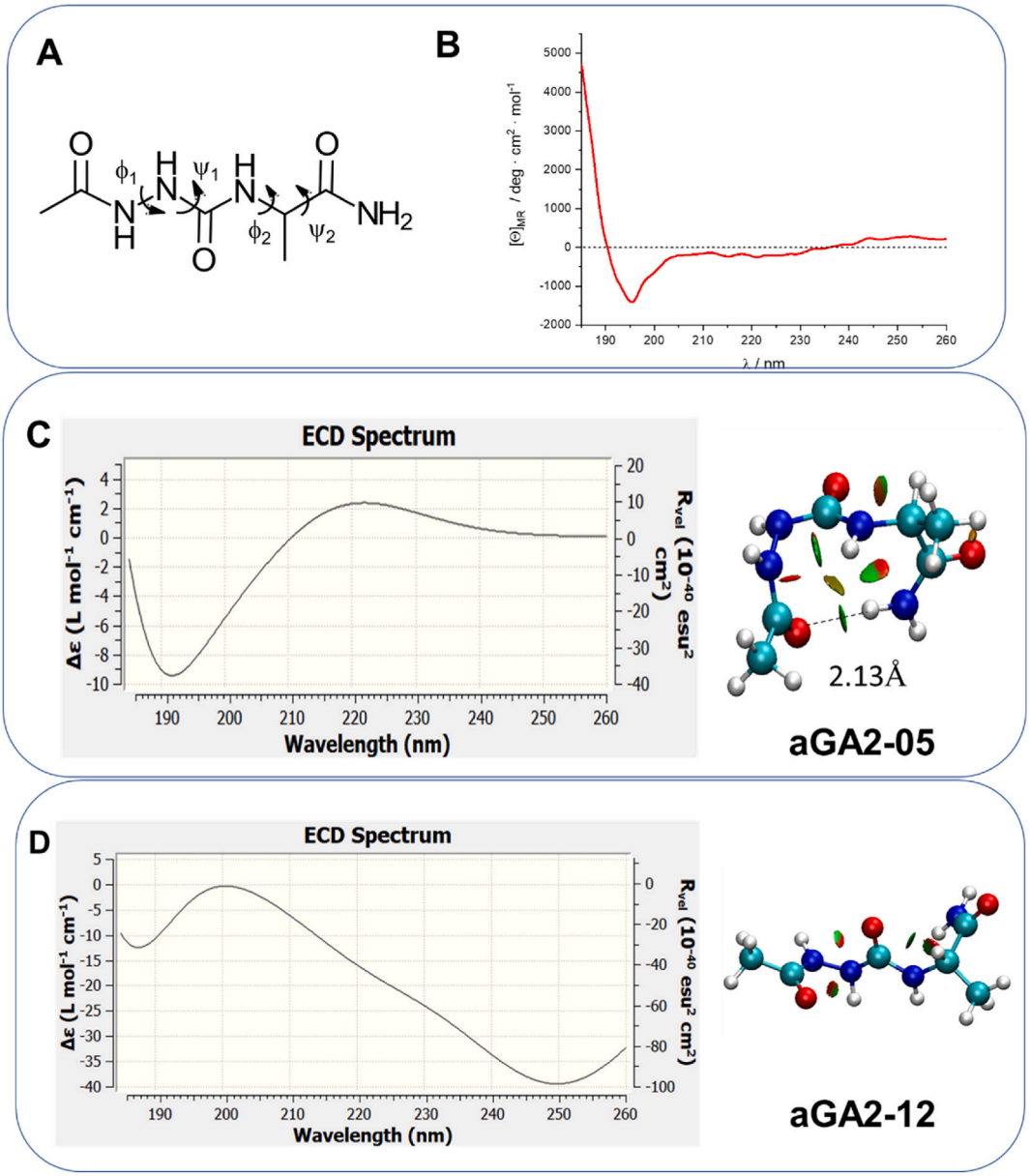


Fig. 10. CD spectrum of azapeptide **5**. (**A**) Chemical structure of azapeptide **5** (**B**) Experimental CD spectrum of azapeptide **5** in water at 5 °C. (**C** and **D**) Calculated electronic circular dichroism (ECD) for **aGA2-05** and **aGA2-12** at the SMD/B3LYP-D3 functional with a 6–311++G(2d,2p) basis set in water. NCI analysis of **aGA2-05** and **aGA2-12**. NCI plot shows the regions of attractive (blue) and repulsive (red) interactions, as well as regions of van der Waals interactions (green) for the B3LYP-D3 functional (Multiwfn and VMD software were used to generate the figures). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

found to be the second lowest energy minimum (ΔE = 0.62 kcal/mol) (SI Tables S15 and S16 and Fig. S4). We observed that SMD/LCgau-BOP + LRD functional predicts the four β–II–turn structures (**FaG02**, **FaG01**, **FaG11**, and **FaG16**) with the different orientation of χ_1 value, corresponding to 55 % of the total population (SI Table S16). However, the SMD/B3LYP-D3 method predicts the conformer **FaG14** (β-I turn, 24 % population) as the lowest energy conformer; the conformer **FaG02** (β-II turn, 14 % population) is the second lowest energy minimum (SI Tables S17 and S18). The results suggest that the SMD/LCgau-BOP + LRD method would explain the population of azapeptide **6** better than the SMD/B3LYP-D3 method in water.

4. Conclusion

The neighbor and solvation effects on the conformational spaces of Ala residues were examined using two azapeptides, Ac-Ala-

azaGly-NHMe (**3**) and Ac-azaGly-Ala-NHMe (**4**). Since azaGly can adopt specific conformation, we could reduce the possible conformations of azapeptide models. Using the LCgau-BOP without and with LRD, we generated the potential energy maps of i+1 or i+2 Ala residue, showing that the local minima of Ala residues are found in the beta-sheet (β_S), inverse gamma-turn(γ'), polyproline II (β_P), the right-handed alpha-helical/bridged (α_R/δ_R) regions depending on the neighbor and solvation effect. The polyproline II conformer in the i+1 position of Ala residue is found as a local minimum as the i+2 azaGly adopts the left-bridged (δ_L: $\varphi = +90°$, $\psi = 0°$) conformation in the gas phase and water. However, the β_P conformer of i+1 Ala residue was found to be local minima regardless of the conformations of azaGly residue in water. The conformers α_R/δ_R were found as local minima for the i+1 or i+2 Ala residue as azaGly adopts the δ_R, β_P, or ε conformer regardless of the sequence and solvent. The IEF-PCM and SMD solvation methods exhibit that the γ' conformer is destabilized in water, although it is stabilized in the gas phase. The X-ray structure of Ac-azaGly-Ala-NH_2 (**5**, **aGA2**) was obtained, showing that the i+1 azaGly residue adopts the β_P conformer and the i+2 Ala adopts the δ_R conformer in the solid state. The β_P conformer of the i+1 azaGly of azapeptide **5** would be stabilized by intermolecular hydrogen bonds [43,76,78], and the δ_R conformer of i+2 Ala residue may be stabilized by solvent. In water, azapeptide **5** adopts β-I turn ($\delta_R\delta_R$ conformer) as shown in the experimental and theoretical CD spectrum of **aGA2-05** conformer. We also assessed the conformational preferences of Ac-Phe-azaGly-NH_2 (**6**, **FaG**) using DFT functionals, showing that SMD/LCgau-BOP + LRD predictions agree well with the experimental results. The results imply that the inclusion of the dispersion effect is necessary to describe the conformational properties of Ala residue in short azapeptides. The study may also be employed to design a new foldamer containing azaglycine residue in peptide.

Data availability statement

The PDB files of the optimized structures of azapeptides **3**–**6** at the different DFT functionals in the gas phase and water are also available in Figshare (https://doi.org/10.6084/m9.figshare.21552297.v1). CCDC **2290581** contains the supplementary crystallographic data for this paper. These data are provided free of charge by the joint Cambridge Crystallographic Data Centre and Fachinformationszentrum Karlsruhe Access Structures service www.ccdc.cam.ac.uk/structures.

CRediT authorship contribution statement

Ho-Jin Lee: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Investigation, Data curation, Conceptualization. **Shi-Wei Liu:** Visualization, Data curation. **Máté Sulyok-Eiler:** Visualization, Data curation. **Veronika Harmat:** Visualization, Data curation. **Viktor Farkas:** Visualization, Data curation. **Zoltán Bánóczi:** Writing – review & editing, Funding acquisition, Data curation. **Mouna El Khabchi:** Visualization, Validation, Data curation. **Hua-Jun Shawn Fan:** Writing – review & editing, Supervision, Funding acquisition. **Kimihiko Hirao:** Writing – review & editing, Software. **Jong-Won Song:** Writing – review & editing, Software, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is dedicated to Dr. Kang-Bong Lee, who retired after a long distinguished scientific expertise and mentorship. This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2023-00278134). HJS Fan would like to acknowledge the partial financial support from the Science Foundation of Sichuan University of Science & Engineering (2020RC06) and the Natural Science Foundation of Sichuan Province (SYZ202133 & 2022JDGD0041). J.-W.S. would like to acknowledge the partial financial support from a Daegu Research Grant, 2023-0314. The numerical calculations were partly conducted on the Hokusai supercomputing facility of RIKEN Information Systems Division (J.-W.S.). The crystallographic study was conducted as part of project No. VEKOP-2.3.2-16-2017-00014 and VEKOP-2.3.3-15-2017-00018 co-funded by the European Union, the State of Hungary, and the European Regional Development Fund, as well as by project No. 2018–1.2.1-NKP-2018-00005 funded by the Hungarian National Research, Development and Innovation Office, Hungary.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33159.

References

[1] A. Perczel, J.G. Angyan, M. Kajtar, W. Viviani, J.L. Rivail, J.F. Marcoccia, I.G. Csizmadia, Peptide models. 1. Topology of selected peptide conformational potential energy surfaces (glycine and alanine derivatives), J. Am. Chem. Soc. 113 (16) (1991) 6256–6265.

[2] T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III, Pople JA: theoretical study of blocked glycine and alanine peptide analogs, J. Am. Chem. Soc. 113 (16) (1991) 5989–5997.
[3] A. Perczel, O. Farkas, I. Jákli, I.A. Topol, I.G. Csizmadia, Peptide models XXXIII. Extrapolation of low-level Hartree-Fock data of peptide conformation to large basis set SCF, MP2, DFT, and CCSD(T) results. The Ramachandran surface of alanine dipeptide computed at various levels of theory, J. Comput. Chem. 24 (9) (2003) 1026–1042.
[4] I. Hudáky, P. Hudáky, A. Perczel, Solvation model induced structural changes in peptides. A quantum chemical study on Ramachandran surfaces and conformers of alanine diamide using the polarizable continuum model, J. Comput. Chem. 25 (12) (2004) 1522–1531.
[5] Z.X. Wang, Y. Duan, Solvation effects on alanine dipeptide: a MP2/cc-pVTZ//MP2/6-31G** study of (Phi, Psi) energy maps and conformers in the gas phase, ether, and water, J. Comput. Chem. 25 (14) (2004) 1699–1716.
[6] C.S. Wang, Y. Zhang, K. Gao, Z.Z. Yang, A new scheme for determining the intramolecular seven-membered ring N-H.O=C hydrogen-bonding energies of glycine and alanine peptides, J. Chem. Phys. 123 (2) (2005) 24307.
[7] V.M. Dadarlat, Potentials of mean force for the interaction of blocked alanine dipeptide molecules in water and gas phase from MD simulations, Biophys. J. 89 (3) (2005) 1433–1445.
[8] Y.K. Kang, Conformational preferences of non-prolyl and prolyl residues, J. Phys. Chem. B 110 (42) (2006) 21338–21348.
[9] P. Echenique, I. Calvo, J.L. Alonso, Quantum mechanical calculation of the effects of stiff and rigid constraints in the conformational equilibrium of the alanine dipeptide, J. Comput. Chem. 27 (14) (2006) 1733–1747.
[10] J. Vymetal, J. Vondrasek, Metadynamics as a tool for mapping the conformational and free-energy space of peptides–the alanine dipeptide case study, J. Phys. Chem. B 114 (16) (2010) 5632–5642.
[11] R. Ishizuka, G.A. Huber, J.A. McCammon, Solvation effect on the conformations of alanine dipeptide: integral equation approach, J. Phys. Chem. Lett. 1 (15) (2010) 2279–2283.
[12] F.F. Garcia-Prieto, I. Fdez Galvan, M.A. Aguilar, M.E. Martin, Study on the conformational equilibrium of the alanine dipeptide in water solution by using the averaged solvent electrostatic potential from molecular dynamics methodology, J. Chem. Phys. 135 (19) (2011) 194502.
[13] Y. Yonezawa, I. Fukuda, N. Kamiya, H. Shimoyama, H. Nakamura, Free energy landscapes of alanine dipeptide in explicit water reproduced by the force-switching wolf method, J. Chem. Theor. Comput. 7 (5) (2011) 1484–1493.
[14] K.M. Ng, M. Solayappan, K.L. Poh, Global energy minimization of alanine dipeptide via barrier function methods, Comput. Biol. Chem. 35 (1) (2011) 19–23.
[15] V. Cruz, J. Ramos, J. Martínez-Salazar, Water-mediated conformations of the alanine dipeptide as revealed by distributed umbrella sampling simulations, quantum mechanics based calculations, and experimental data, J. Phys. Chem. B 115 (16) (2011) 4880–4886.
[16] S. Toal, R. Schweitzer-Stenner, Local order in the unfolded state: conformational biases and nearest neighbor interactions, Biomolecules 4 (3) (2014) 725–773.
[17] J. Rubio-Martinez, M.S. Tomas, J.J. Perez, Effect of the solvent on the conformational behavior of the alanine dipeptide deduced from MD simulations, J. Mol. Graph. Model. 78 (2017) 118–128.
[18] G.A. Balaji, H.G. Nagendra, V.N. Balaji, S.N. Rao, Experimental conformational energy maps of proteins and peptides, Proteins 85 (6) (2017) 979–1001.
[19] V. Mironov, Y. Alexeev, V.K. Mulligan, D.G. Fedorov, A systematic study of minima in alanine dipeptide, J. Comput. Chem. 40 (2) (2019) 297–309.
[20] S. Dasari, B.S. Mallik, Conformational free-energy landscapes of alanine dipeptide in hydrated ionic liquids from enhanced sampling methods, J. Phys. Chem. B 124 (31) (2020) 6728–6737.
[21] C. Cabezas, M. Varela, V. Cortijo, A.I. Jimenez, I. Pena, A.M. Daly, J.C. Lopez, C. Cativiela, J.L. Alonso, The alanine model dipeptide Ac-Ala-NH_2 exists as a mixture of C_7(eq) and C_5 conformers, Phys. Chem. Chem. Phys. 15 (7) (2013) 2580–2585.
[22] A. Mirtic, F. Merzel, J. Grdadolnik, The amide III vibrational circular dichroism band as a probe to detect conformational preferences of alanine dipeptide in water, Biopolymers 101 (7) (2014) 814–818.
[23] D. Caballero, J. Määttä, A.Q. Zhou, M. Sammalkorpi, C.S. O'Hern, L. Regan, Intrinsic α-helical and β-sheet conformational preferences: a computational case study of alanine, Protein Sci. 23 (7) (2014) 970–980.
[24] R.A. Cormanich, M. Bühl, R. Rittner, Understanding the conformational behaviour of Ac-Ala-NHMe in different media. A joint NMR and DFT study, Org. Biomol. Chem. 13 (35) (2015) 9206–9213.
[25] F. Jiang, W. Han, Y.-D. Wu, The intrinsic conformational features of amino acids from a protein coil library and their applications in force field development, Phys. Chem. Chem. Phys. 15 (10) (2013) 3413–3428.
[26] D. Wei, H. Guo, D.R. Salahub, Conformational dynamics of an alanine dipeptide analog: an ab initio molecular dynamics study, Phys. Rev. E - Stat. Nonlinear Soft Matter Phys. 64 (1 Pt 1) (2001) 011907.
[27] H. Hu, M. Elstner, J. Hermans, Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution, Proteins 50 (3) (2003) 451–463.
[28] J. Grdadolnik, V. Mohacek-Grosev, R.L. Baldwin, F. Avbelj, Populations of the three major backbone conformations in 19 amino acid dipeptides, Proc. Natl. Acad. Sci. U.S.A. 108 (5) (2011) 1794–1798.
[29] V. Parchansky, J. Kapitan, J. Kaminsky, J. Sebestik, P. Bour, Ramachandran plot for alanine dipeptide as determined from Raman optical activity, J. Phys. Chem. Lett. 4 (16) (2013) 2763–2768.
[30] M. Culka, T. Kalvoda, O. Gutten, L. Rulíšek, Mapping conformational space of all 8000 tripeptides by quantum chemical methods: what strain is affordable within folded protein chains? J. Phys. Chem. B 125 (1) (2021) 58–69.
[31] T. Kalvoda, M. Culka, L. Rulíšek, E. Andris, Exhaustive mapping of the conformational space of natural dipeptides by the DFT-D3//COSMO-RS method, J. Phys. Chem. B 126 (32) (2022) 5949–5958.
[32] B. Chahkandi, M. Chahkandi, A reconnaissance DFT study of the full conformational analysis of N−formyl−L−serine−L−alanine−NH_2 dipeptide, J. Mol. Model. 26 (6) (2020).
[33] V.K. Prasad, A. Otero-de-la-Roza, G.A. DiLabio, PEPCONF, a diverse data set of peptide conformational energies, Sci. Data 6 (2019) 180310.
[34] P.A. Karplus, Experimentally observed conformation-dependent geometry and hidden strain in proteins, Protein Sci. 5 (7) (1996) 1406–1420.
[35] K. Fan Cheng, S. VanPatten, M. He, Y. Al-Abed, Azapeptides -A history of synthetic milestones and key Examples, Curr. Med. Chem. 29 (42) (2022) 6336–6358.
[36] J. Wątły, A. Miller, H. Kozłowski, M. Rowińska-Żyrek, Peptidomimetics - an infinite reservoir of metal binding motifs in metabolically stable and biologically active molecules, J. Inorg. Biochem. 217 (2021) 111386.
[37] A. Begum, D. Sujatha, K.V.S.R.G. Prasad, K. Bharathi, A Review on azapeptides: the promising peptidomimetics, Asian J. Chem. 29 (9) (2017) 1879–1887.
[38] M. El Khabchi, H. Lahlou, Z. El Adnani, M. McHarfi, M. Benzakour, A. Fitri, A.T. Benjelloun, Conformational preferences of Ac-Pro-azaXaa-NHMe (Xaa = Asn, Asp, Ala) and the effect of intramolecular hydrogen bonds on their stability in gas phase and solution, J. Mol. Model. 27 (12) (2021) 368.
[39] Khabchi M. El, M. Mcharfi, M. Benzakour, A. Fitri, A.T. Benjelloun, J.W. Song, K.B. Lee, H.J. Lee, Computational investigation of conformational properties of short azapeptides: Insights from DFT study and NBO analysis, Molecules 28 (14) (2023) 5454.
[40] K. Tarchoun, D. Soltész, V. Farkas, H.J. Lee, I. Szabó, Z. Bánóczi, Influence of Aza-Glycine substitution on the internalization of penetratin, Pharmaceutics 16 (4) (2024) 477.
[41] H.J. Lee, J.W. Song, Y.S. Choi, H.M. Park, K.B. Lee, A Theoretical study of conformational properties of N-methyl azapeptide derivatives, J. Am. Chem. Soc. 124 (40) (2002) 11881–11893.
[42] H.J. Lee, J.W. Song, Y.S. Choi, S. Ro, C.J. Yoon, The Energetically favorable cis peptide bond for the azaglycine-containing peptide: for-AzGly-NH_2 model, Phys. Chem. Chem. Phys. 3 (9) (2001) 1693–1698.
[43] S.D. Melton, E.A.E. Brackhahn, S.J. Orlin, P. Jin, D.M. Chenoweth, Rules for the design of aza-glycine stabilized triple-helical collagen peptides, Chem. Sci. 11 (39) (2020) 10638–10646.
[44] T. Harris, D.M. Chenoweth, Sterics and Stereoelectronics in Aza-Glycine: impact of aza-glycine preorganization in triple helical collagen, J. Am. Chem. Soc. 141 (45) (2019) 18021–18029.

[45] F.A. Etzkorn, R.I. Ware, A.M. Pester, D. Troya, Conformational analysis of n→π* interactions in collagen triple helix models, J. Phys. Chem. B 123 (2) (2019) 496–503.
[46] S.D. Melton, D.M. Chenoweth, Variation in the Yaa position of collagen peptides containing azaGlycine, Chem. Commun. 54 (84) (2018) 11937–11940.
[47] H.J. Lee, H.J. Jung, J.H. Kim, H.-M. Park, K.B. Lee, Conformational preference of azaglycine-containing dipeptides studied by PCM and IPCM methods, Chem. Phys. 294 (2) (2003) 201–210.
[48] S. Ro, H.J. Lee, I.A. Ahn, D.K. Shin, K.B. Lee, C.J. Yoon, Y.S. Choi, Torsion angle based design of peptidomimetics: a dipeptidic template adopting β-I Turn (Ac-Aib-AzGly–NH_2), Bioorg. Med. Chem. 9 (7) (2001) 1837–1841.
[49] X.S. Yan, H. Luo, K.S. Zou, J.L. Cao, Z. Li, Y.B. Jiang, Short azapeptides of folded structures in aqueous solutions, ACS Omega 3 (5) (2018) 4786–4790.
[50] J.-W. Song, T. Hirosawa, T. Tsuneda, K. Hirao, Long-range corrected density functional calculations of chemical reactions: redetermination of parameter, J. Chem. Phys. 126 (15) (2007) 154105.
[51] T. Sato, H. Nakai, Local response dispersion method. II. Generalized multicenter interactions, J. Chem. Phys. 133 (19) (2010) 194101.
[52] T. Sato, H. Nakai, Density functional method including weak interactions: dispersion coefficients based on the local response approximation, J. Chem. Phys. 131 (22) (2009) 224104.
[53] Y. Ikabata, H. Nakai, Local response dispersion method: a density-dependent dispersion correction for density functional theory, Int. J. Quant. Chem. 115 (5) (2015) 309–324.
[54] J.-W. Song, T. Tsuneda, T. Sato, K. Hirao, An examination of density functional theories on isomerization energy calculations of organic molecules, Theor. Chem. Acc. 130 (4) (2011) 851–857.
[55] J.-W. Song, T. Tsuneda, T. Sato, K. Hirao, Calculations of alkane energies using long-range corrected DFT combined with intramolecular van der Waals correlation, Org. Lett. 12 (7) (2010) 1440–1443.
[56] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (15) (2010) 154104.
[57] M.J. Frisch Gwt, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, 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, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman, D.J. Fox, Gaussian 09, Revision A.02, Gaussian, Inc., 2010.
[58] M.J. Frisch Gwt, H.B. Schlegel, G.E. Scuseria, M.A. Robb Jrc, G. Scalmani, V. Barone, G.A. Petersson Hn, X. Li, M. Caricato, A.V. Marenich, J. Bloino Bgj, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz Afi, J.L. Sonnenberg, D. Williams-Young, F. Ding Fl, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson Dr, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng Wl, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa Mi, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven Kt, J.A. Montgomery Jr., J.E. Peralta, et al., Gaussian 16, Revision C.01, Gaussian, Inc., 2019.
[59] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys Rev A Gen Phys 38 (6) (1988) 3098–3100.
[60] T. Tsuneda, T. Suzumura, K. Hirao, A new one-parameter progressive Colle–Salvetti-type correlation functional, J. Chem. Phys. 110 (22) (1999) 10664–10678.
[61] J.-W. Song, S. Tokura, T. Sato, M.A. Watson, K. Hirao, An improved long-range corrected hybrid exchange-correlation functional including a short-range Gaussian attenuation (LCgau-BOP), J. Chem. Phys. 127 (15) (2007) 154109.
[62] R. Kar, J.W. Song, T. Sato, K. Hirao, Long-range corrected density functionals combined with local response dispersion: a promising method for weak interactions, J. Comput. Chem. 34 (27) (2013) 2353–2359.
[63] G. Scalmani, M.J. Frisch, Continuous surface charge polarizable continuum models of solvation. I. General formalism, J. Chem. Phys. 132 (11) (2010) 114110.
[64] A.V. Marenich, C.J. Cramer, D.G. Truhlar, Universal Solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions, J. Phys. Chem. B 113 (18) (2009) 6378–6396.
[65] M. Froimowitz, HyperChem: a software package for computational chemistry and molecular modeling, Biotechniques 14 (6) (1993) 1010–1013.
[66] N. Lima, G. Rocha, R. Freire, A. Simas, RM1 semiempirical model: chemistry, pharmaceutical research, molecular biology and materials science, J. Braz. Chem. Soc. 30 (4) (2019) 683–716.
[67] R.K. Dennington, A. Todd, John M. Millam, GaussView, Version 6, 2016.
[68] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunction analyzer, J. Comput. Chem. 33 (5) (2012) 580–592.
[69] Z. Bánóczi, Á. Tantos, A. Farkas, Z. Majer, L.E. Dókus, P. Tompa, F. Hudecz, New m-calpain substrate-based azapeptide inhibitors, J. Pept. Sci. 19 (6) (2013) 370–376.
[70] O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, OLEX2: a complete structure solution, refinement and analysis program, J. Appl. Crystallogr. 42 (2) (2009) 339–341.
[71] G.M. Sheldrick, SHELXT– Integrated space-group and crystal-structure determination, Acta Crystallographica Section A Foundations and Advances 71 (1) (2015) 3–8.
[72] G.M. Sheldrick, Crystal structure refinement with SHELXL, Acta Crystallogr., Sect. C: Struct. Chem. 71 (1) (2015) 3–8.
[73] P.R. Edgington, P. McCabe, C.F. Macrae, E. Pidcock, G.P. Shields, R. Taylor, M. Towler, J. Van De Streek, Mercury: visualization and analysis of crystal structures, J. Appl. Crystallogr. 39 (3) (2006) 453–457.
[74] A.L. Spek, Structure validation in chemical crystallography, Acta Crystallogr D Biol Crystallogr 65 (Pt 2) (2009) 148–155.
[75] Y. Zhang, R.M. Malamakal, D.M. Chenoweth, Aza-glycine induces collagen hyperstability, J. Am. Chem. Soc. 137 (39) (2015) 12422–12425.
[76] A.J. Kasznel, Y. Zhang, Y. Hai, D.M. Chenoweth, Structural basis for Aza-glycine stabilization of collagen, J. Am. Chem. Soc. 139 (28) (2017) 9427–9430.
[77] T. Kastinen, D. Lupa, P. Bonarek, D. Fedorov, M. Morga, M.B. Linder, J.L. Lutkenhaus, P. Batys, M. Sammalkorpi, pH dependence of the assembly mechanism and properties of poly(L-lysine) and poly(L-glutamic acid) complexes, Phys. Chem. Chem. Phys. 25 (27) (2023) 18182–18196.
[78] S.D. Melton, M.S. Smith, D.M. Chenoweth, Incorporation of aza-glycine into collagen peptides, J. Org. Chem. 85 (3) (2020) 1706–1711.