Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

D-Amino acid substituted peptides as potential alternatives of homochiral L-configurations

Jianxun Shen [1,*]

[1] School of Earth and Environmental Sciences and Centre for Exoplanet Science, University of St Andrews, St Andrews KY16 9AL, UK

* Correspondence: js365@st-andrews.ac.uk

Abstract

On the primitive Earth, both L- and D-amino acids would have been present. However, only L-amino acids are essential blocks to construct proteins in modern life. To study the relative stability of D-amino acid substituted peptides, a variety of computational methods were applied. Ten prebiotic amino acids (Gly, Ala, Asp, Glu, Ile, Leu, Pro, Ser, Thr, and Val) were previously determined by multiple meteorite, spark discharge, and hydrothermal vent studies. Some previously reported early Earth polypeptide analogs were focused on in this study. Tripeptides composed of only Asp, Ser, and Val exemplified that different positions (i.e., N-terminus, C-terminus, and middle) made a difference in the minimal folding energy of peptides, while the chemical classification of amino acid (hydrophobic, acidic, or hydroxylic) did not show a significant difference. Hierarchical cluster analysis for dipeptides with all possible combinations of the proposed ten prebiotic amino acids and their D-amino acid substituted derivatives generated five clusters. Primordial simple polypeptides were modeled to test the significance of molecular fluctuations, secondary structure occupancies, and folding energy differences based on these clusters. We found peptides with α-helices, long β-sheets, and long loops are usually less sensitive to D-amino acid replacements in comparison to short β-sheets. Intriguingly, amongst 129 D-amino acid residues, mutation sensitivity profiles presented that the ratio of more to less stable residues was about 1. In conclusion, some combinations of a mixture of L- and D-amino acids can potentially act as essential building blocks of life.

Keywords: D-amino acids; homochirality; molecular modeling; secondary structure occupancy; stability

1 INTRODUCTION

The scientific question we attempt to address and explore in this study is that whether homochirality is the necessary biosignature of all organisms. Homochirality means that the enantiomeric excess is approximately 100% in an enantiomer mixture. Within life systems, all protein-derived amino acids in organisms are in the L-configuration, and all monomers of polysaccharides are in the D-configuration, with the exception of few prokaryotes (Morneau et al. 2014). The other configuration of biomolecules usually causes damages to the health of life (Paulson and Shug 1981). In effect, the evolution of biomolecules tends to minimize folding energy to stabilize their structures, as a typical reflection of the second law of thermodynamics. However, the chemical self-replicative reactions in biological systems (Buschmann et al. 2000; Wu et al. 2012), such as homochirality, are far from the favorable thermodynamic equilibrium, since thermodynamic equilibrium prefers racemic mixture (Pross 2012). It is crucial to investigate how primitive chemical blocks are built up for macromolecules and further for complex organisms.

Curiosity about the hypothesis of primordial prebiotic amino acid species expands after 1952's Miller-Urey experiment. These early Earth spark discharge simulations consistently yielded 10 abundant racemic individual amino acids, including glycine (Gly or G), alanine (Ala or A), aspartic acid (Asp or D),

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

glutamic acid (Glu or E), isoleucine (Ile or I), leucine (Leu or L), proline (Pro or P), serine (Ser or S), threonine (Thr or T), and valine (Val or V) (Bada and Miller 1987; Miller 1953; Miller 1974; Ring et al. 1972; Zaia et al. 2008), identical to various studies on extraterrestrial comets or meteorites and hydrothermal vents (Longo and Blaber 2012). Hereafter, Gly, Ala, Val, Pro, Ile, and Leu are collectively termed as prebiotic hydrophobic amino acids; Asp and Glu are collectively termed as prebiotic acidic amino acids; and Ser and Thr are collectively termed as prebiotic hydroxylic amino acids. Among these amino acids, Gly is the only achiral amino acid, and the other 9 are chiral amino acids. Both L- and D-configurations of 9 chiral prebiotic amino acids have been observed in natural peptides, and natural D-amino acids were mostly discovered in short antibiotic or immunosuppressive peptides, such as gramicidin, actinomycin, valinomycin, tyrocidine, and cyclosporine (Mitchell and Smith 2003; Morneau et al. 2014; Paulson and Shug 1981; Pross 2012; Soda and Osumi 1969). In the modern world, most natural peptides with D-amino acids are usually shorter than 20 amino acids (Mitchell and Smith 2003) and D-amino acids seemingly play little role in biological systems.

Admissions should be made that D-amino acid substitutions of L-peptides often resulted in the deconstructions of protein conformations (Krause et al. 1995; Nanda and Degrado 2004; Zerze et al. 2018), and homochirality was a more efficient way to store and retrieve spatial information (Carroll 2009). However, D-amino acids sometimes were beneficial to organismic health. For instance, evidence supported that D-serine and D-aspartate could modulate and mitigate neurodegradation, schizophrenic symptoms, epilepsy, ischemia, and amyotrophic lateral sclerosis (Billard 2012; Fuchs et al. 2005; Wolosker 2011). D-arginine protected the central nervous system from neurotoxicity and avoided the onset of mental disorders without noticeable side effects (Canteros 2014). Although the function of D-amino acid oxidase was to prevent the harmful accumulation of D-amino acids (Daniello et al. 1993), knockouts of D-amino acids oxidase, which inhibited the degradation of D-amino acids, could improve spatial cognition, odor cognition, and short-term spontaneous recognition memory performance for several days (Pritchett et al. 2015). Moreover, various amino acid racemases that transform L-amino acids to D-amino acids exist in nature (Soda and Osumi 1969). Skolnick et al. (2019) applied computational modeling of D-amino acid substitutions and found that peptides composed of a mixture of L- and D-amino acid residues could be involved in ancient metabolisms in the prebiotic world despite less thermodynamically stable structures (Skolnick et al. 2019). As above, it is rational to admit that peptides with D-amino acids are potential to replace this homochiral amino acid living world.

As the computational modeling technology progresses, we have more powerful tools to study peptide structures and folding processes. The strategy to stabilize D-amino acids in a sequence primarily constituted by L-amino acids was to minimize the χ-angle of peptide bonds and the placement of amino acid backbone for an overall minimal steric hindrance. In the open-source package suite Rosetta, the talaris13 force field can be used for scoring not only canonical L-amino acids but also their D-amino acid mirror coordinates (O'Meara et al. 2015). D-amino acid rotamer libraries for the rotation and flexibility predictions were introduced in (Renfrew et al. 2012) and (Garton et al. 2017). The force field of FoldX performs slightly better than other programs in predicting stability energies (Garton et al. 2017). Besides, the Optimized Potentials for Liquid Simulations 3 (OPLS3) force field has a great accuracy in modeling small peptides such as dipeptides (Storer et al. 1995).

Given the importance of homochirality, in this study, we homochirally transformed targeted L-amino acid species in primordial simple peptide analogs (hereafter primordial peptides) into their respective D-configurations and investigated these new peptides after energy minimization. The purpose of this study is to analyze whether a mixture of homochiral L- and D-amino acid residues could stabilize the folding energy and sustain the folding structure. We analyzed and compared the structure and stability of each peptide and D-amino acid substituted derivatives in terms of a variety of aspects, including secondary

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

structure types, packing energy, B-factor (the fluctuation of structure caused by thermal motion), attractive force, repulsive force, solvation energy, electrostatic potential, hydrogen bonds (H-bonds), and disulfide geometry potential. Statistical analyses were further performed to explore the mutuality between these variables. Our results implicate that some combinations of mixed homochiral L- and D-amino acids can act as essential building blocks for biological systems beyond Earth.

2 MATERIALS AND METHODS

2.1 Oligopeptide modeling and analyses

To investigate the similarity of prebiotic D-amino acids with respect to how they affected stability energy (the change in Gibbs free energy of the product after mutation from the starting material during the energy minimization process (Potapov et al. 2009; Thiltgen and Goldstein 2012)), we modeled 99 dipeptides composed of all combinations of prebiotic L-amino acids (Gly, Ala, Asp, Glu, Ile, Leu, Pro, Ser, Thr, and Val) except for Gly-Gly since it was the only achiral dipeptide. In addition, 180 monosubstituted D-amino acid derivatives of these dipeptides were modeled using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). The energy minimization of dipeptides was performed using the OPLS3 force field with aqueous solvation (Harder et al. 2016; Shivakumar et al. 2012; Storer et al. 1995) by Schrödinger Maestro 11 (Schrödinger Release 2019-4: Maestro, Schrödinger, LLC, New York, NY, 2019.). The stability energy, in the unit of kilocalorie/mole (kcal/mol), of dipeptides was determined by FoldX (Schymkowitz et al. 2005). The stability energy differences of monosubstituted D-amino acid derivatives from that of their correspondent L-dipeptides were calculated. Based on these 180 sets of stability difference data, 10 prebiotic D-amino acids were statistically grouped by hierarchical clustering analysis (using unweighted pair group method with arithmetic mean [UPGMA] algorithm and Euclidean distance matrix) in Past 4.03 (Hammer et al. 2001).

To study whether the chemical classification and the position of residue significantly influenced stability energy, tripeptides composed of Val, Asp, and Ser were modeled, since these 3 amino acids represented abundant prebiotic hydrophobic, acidic, and hydroxylic amino acids, respectively. In total, 6 L-dipeptides (i.e., DSV, DVS, SDV, SVD, VDS, and VSD) and their 42 mono- and polysubstituted D-amino acid derivatives were built. After energy minimization and stability determination, the differences of stability energy after substitutions were calculated. Then, Kruskal-Wallis independent sample tests were performed to examine the significance of the effect of amino acid species and positions on stability differences using IBM SPSS Statistics 25 (IBM Corp. 2019).

2.2 D-amino acid substitutions of polypeptides

Herein, polypeptides substituted with one or more D-amino acids were designated using lowercase letters, followed by the number and species of substitutions. For example, 1ark_9dA is the D-mutant of the all L-amino acid parent polypeptide 1ARK that switched all 9 L-Ala residues to D-Ala. For polypeptides substituted by multiple species of D-amino acids, the "d" in front of D-residue names would be omitted to save space, such as 1ark_9A1P6D3E5I4V4T. Based on a preliminary study on several randomly D-substituted polypeptides, we found they did not have predictable stable secondary structures and stability energies, in agreement with a previous study (Rodriguez-Granillo et al. 2011). Therefore, we homochirally, instead of randomly, transformed the chosen prebiotic L-amino acids into their D-configurations in this study. Please note that modeling and structure predictions applied in this study were not *ab initio* methods from unfolded amino acid chains. Instead, the structure of original

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

homochiral L-amino acid peptides were used as a starting point, and additional modifications (i.e., chirality inverting and energy minimization) were performed based on the original structure.

To study the D-amino acid substitutions in longer peptides, several primordial polypeptides were selected from Minervini et al. (2015) and Chao et al. (2012). Maestro 11 was used for the chirality transformations of polypeptides. The D-amino acid rotamer library was incorporated into Rosetta package by following (Renfrew et al. 2012) and (Garton et al. 2017). Monosubstituted D-amino acid polypeptides of all primordial peptides proposed by Minervini et al. (2015) (i.e., 1ARK, 1HX2, 1KV0, 1PPT, 1TCP, 1UOY, 1ZFI, 1BHI, 2BHI, 2KJF, 2CDX) (Minervini et al. 2015) and 2LZE (Chao et al. 2013) were modeled. The pose of residue geometry was mirrored using the Rosetta FlipChirality mover. Torsion and Cartesian energy minimizations were performed using the talaris13 force field with aqueous solvation (O'Meara et al. 2015) by PyRosetta3 (Chaudhury et al. 2010) and InteractiveROSETTA 2.3 (Schenkelberg and Bystroff 2015). Root mean square deviation (RMSD) calculations of polypeptides were performed in InteractiveROSETTA 2.3. B-factors and the occupancies of secondary structures, including α-helix, β-sheet, β-turn, and 3-10 helix, were measured in Yet Another Scientific Artificial Reality Application (YASARA) (Krieger and Vriend 2014). Force field of the FoldX was applied for the calculation of the stability energy. The bivariate correlation and multiple regression analyses between stability energy, secondary structure occupancies, B-factors, and the number of D-mutations were performed using IBM SPSS Statistics 25.

Amongst the aforementioned 12 polypeptides, 4 of them (1ARK, 1PPT, 1ZFI, and 2LZE) (Table 1) were nominated for further analyses. The reason to focus on these 4 polypeptides was that they possess all 10 proposed prebiotic amino acids and represent different types of secondary structures: in accordance with PyMOL visualizations, 1ARK is characterized by 8 dominant anti-parallel short β-sheets; 1PPT is characterized by 1 dominant long α-helix; 1ZFI is characterized by 5 anti-parallel β-sheets, 1 short α-helix, and 4 disulfide bonds; and 2LZE is characterized by 2 dominant zinc loop arms and 1 short fragment of 3-10 helix (Table 2). According to the prebiotic amino acid clusters grouped based on D-amino acid substituted dipeptides, the L-configurations of these amino acid clusters in 1ARK, 1PPT, and 1ZFI were substituted for all combinations of their D-counterparts to avoid excessive combinative polysubstituted peptide models. These D-cluster substituted models were folded using Torsion and Cartesian energy minimization methods. RMSDs and stability energies were determined by InteractiveROSETTA 2.3 and FoldX, respectively. All D-polypeptides 1ark, 1ppt, and 1zfi composed of only D-residues were modeled to examine the symmetry of pose and force field. Additionally, the suggested sets of D-amino acids that could maintain overall structure and improve stability energy were further tested in 1HX2 and 1BHI, for they have all determined amino acid clusters.

Furthermore, the mutation sensitivity profiles of monosubstituted D-amino acid 1ARK, 1PPT, 1ZFI, and 2LZE that illustrated the changes in Gibbs free energy ($\Delta\Delta G$) of each individual D-amino acid after being replaced by all 20 proteinogenic L-amino acids were depicted using the online MAESTROweb (Multi AgEnt STability pRedictiOn) tool (Laimer et al. 2015). In mutation sensitivity profiles, $\Delta\Delta G$ was positively associated with how D-amino acid substituted polypeptide could stabilize the overall structure. Mutants with equal, higher, and lower $\Delta\Delta G$ were then counted.

3 RESULTS AND DISCUSSION

3.1 Explorations of modeled oligopeptides

Prebiotic D-amino acids were clustered to understand the similarity between their effects on changes in stability energy. Cutting at the Euclidean distance of 2 (Figure 1), the hierarchical clustering analysis of D-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

amino acid substituted dipeptides grouped 9 prebiotic D-amino acids into five clusters – (1) Ala and Pro (or AP cluster), (2) Asp (or D cluster), (3) Glu (or E cluster), (4) Leu and Ser (or LS cluster), and (5) Ile, Val, and Thr (or IVT cluster) (Table S1). In all dipeptide models, D-mutations of the AP, D, E, LS, and IVT clusters triggered an increase in energy by -0.057 kcal/mol, +1.17 kcal/mol, +1.42 kcal/mol, +1.42 kcal/mol, and +2.20 kcal/mol, respectively (Table S1). Hereafter, "cluster(s)" refer to these amino acid classification groups determined by hierarchical clustering.

Tripeptide analyses indicated that different positions (i.e., N-terminus, middle, or C-terminus) of D-mutations made a difference in the stability energy of peptides (Kruskal-Wallis test statistic= 6.22*, p= 0.045). This result was consistent with a previous investigation on tripeptide residue positions (Marchesan et al. 2014). However, the chemical property-based classification (acidic, hydroxyl, or hydrophobic) of amino acids did not show a significant difference (Kruskal-Wallis test statistic= 3.59, p= 0.17) without changing positions in these tripeptides (Table S2).

3.2 Correlation and regression analyses of polypeptides

Secondary structure occupancies, B-factors, the number of D-mutations, and stability energies of all 11 primordial polypeptides proposed by Minervini et al. (2015) and their D-mutants were determined and enumerated in Table S3. Automatic linear modeling analysis generated the linear model below:

stability energy = 8.6× number of mutations + 136× coil occupancy - 86× β-sheet occupancy + 1.4

the stability energy was significantly associated with the number of mutations (importance= 0.50***, p< 0.001), coil occupancy (importance= 0.26**, p= 0.002), and β-sheet occupancy (importance= 0.15*, p= 0.02).

Results of the bivariate correlation analyses demonstrated that the stability energy significantly correlated with the number of mutations (Pearson correlation r= 0.41***, p< 0.001) and the occupancy of coils (r= 0.44***, p< 0.001), while significantly anticorrelated with the β-sheet occupancy (r= -0.29**, p= 0.003). Additionally, although not statistically significant, the stability energy was negatively associated with the occupancies of α-helices (r= -0.14, p= 0.09) and β-turns (r= -0.071, p= 0.25), and positively associated with the 3-10 helix occupancy (r= 0.072, p= 0.25) (Table S3).

These outputs indicated that α-helix, β-sheet, and β-turn were capable of stabilizing polypeptide folding, while random loop and 3-10 helix could destabilize the folded structure, which is consistent with our previous structure analyses.

Considering the significant correlation between stability and the number of D-mutations, the effect of D-mutations on stability was hereafter normalized by the following equation:

$$\text{Normalized stability difference} = \frac{\text{Stability of D mutants} - \text{Stability without mutations}}{\text{Number of mutations}}$$

3.3 Secondary structure analyses of polypeptides

The torsion energy minimization method optimized and predicted folding structures without fixing the overall Cartesian coordinates of residues in the original polypeptides using Monte-Carlo algorithm. After flipping the chirality poses of D-residues (Figure S1A, B, and C), all L-polypeptides (1ARK, 1PPT, and 1ZFI) and their corresponding all D-polypeptides had identical stability energies of 109.12 kcal/mol, 20.63 kcal/mol, and 127.41 kcal/mol, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

The 1ARK peptide was relatively unstable because only the D-mutants with three D-Glu (Figure 2Aiv), one D-Leu (Figure 2Avi), or one D-Ser (Figure 2Aviii) were capable of maintaining the original tertiary structure, while the other 6 D-mutations all deconstructed this conformation (Figure 2Aii, Aiii, Av, Avii, Aix, and Ax). The reason for the instability of 1ARK D-mutants attributed to its dominant short β-sheets and the absence of helical secondary structure within its original all L-configuration structure (Ashkenasy et al. 2004; Bourbo et al. 2011; Jackel et al. 2006; Lee et al. 1996; Li and Chmielewski 2003).

As for 1PPT peptide, no matter which species of L-amino acid residue was D-mutated, the dominant α-helix segment did not have a dramatic shape transformation, instead, only a small angular twist at the end of the helical barrel or at the central body part (e.g., the bent α-helix barrel in 1ppt_1dA, Figure 2Bii, as well as that observed previously (Nanda and Degrado 2004)). The D-mutated residues not necessarily affected the shape of their proximal region, such as the D-mutations of D, I, L, T, and V (Figure 2Biii, Bv, Bvi, Bix, and Bx) within the long α-helical structure, while the D-mutation of A affected the original α-helical structure more apparently (Figure 2Bii). Overall, the α-helix was a stable secondary structure to preserve protein conformations (Serrano and Fersht 1989; Yang et al. 1997).

1ZFI had five long anti-parallel β-sheets and one stable helix. However, compared to 1PPT, the helical structure within 1ZFI was much shorter. The long β-sheets in D-amino acid substituted 1ZFI could be divided into several shorter sheets to stabilize the functional structure (e.g., Figure 2Cix).

2LZE was a special primordial polypeptide for its high flexibility of two long loops: 20 different structural models could be found in the Protein Data Bank (PDB) (Burley et al. 2019), and their RMSD values were significantly different from each other. All D-amino acid substituted 2LZE peptides shared similar loop-turn-helix-loop structure with two obvious zinc arms (Figure 2Dii-Dx), indicating the maintenance of its functional domain.

In all cases of the polypeptides examined, proline made a big difference to the original structure because of its rigid ring structure (Yun et al. 1991). Taken all together, peptides with α-helices, long β-sheets, and long loops were less sensitive to D-amino acid substitutions in comparison to those with short β-sheets.

In addition to torsion energy minimization, the Cartesian energy minimization method that concomitantly constrained residues' Cartesian coordinates and optimized peptide stability energy was applied to preserve the functional domains of D-mutant polypeptides. The stability energy differences from their respective original all L-amino acid polypeptides were also calculated. Cartesian minimized results demonstrated that the stability energies of several D-mutant polypeptides (i.e., 1ark_3dE, 1ark_1dS, 1ppt_1dA, 1ppt_1dI, 1ppt_1dS, 1zfi_6dE, 1zfi_4dS, 1zfi_5dT, and 2lze_8dD) were equivalent to their respective all L-amino acid polypeptides (Table 3).

3.4 Hydrogen bonding and hydrophobic interactions

In order to further investigate how the predicted structures of 1ARK, 1PPT, 1ZFI, and 2LZE were formed, hydrogen bonding and hydrophobic interactions that were essential to supersecondary and tertiary structures were visualized in YASARA. The lengths (indicative of strengths) of each hydrogen bonds were listed in Table S4, and bonds formed by D-amino acids were labeled red. Hydrophobic interactions from hydrocarbon chains universally participated in the intramolecular interactions between residues (Figure 3).

1ARK had 8 anti-parallel β-sheets connected through H-bonds. The tertiary structure of 1ARK was folded by multiple H-bonds between anti-parallel β-sheets and multiple hydrophobic interactions throughout its primary sequence (Figure 3Ai). D-amino acid substituted 1ark_6dD, 1ark_3dE, 1ark_1dL, and 1ark_1dS

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

had expanded β-sheets compared to 1ARK (Figure 3Aiii, Aiv, Avi, and Aviii). This finding suggests that D-mutations were possible to produce long β-sheet motifs. However, in 1ark_9dA, 1ark_5dI, 1ark_1dP, 1ark_4dT, and 1ark_4dV, H-bonds only occurred at some turning points of the loop structure because of their less folded structure (Figure 3Aii, Av, Avii, Aix, and Ax). Acidic amino acids (D and E) possessed an O-rich carboxyl sidechain and thus could form H-bonds more flexibly. 1ark_3dE formed one more H-bonds than parent 1ARK. Other D-amino acid monosubstituted 1ark mutants all had less H-bonds, but D-amino acids were actively involved in H bonding (Table S4, 1ARK worksheet tab). Depending on the new folding conformation, the lengths of H-bonds between the same pair of residues varied among mutants.

The α-helix of 1PPT was stabilized by abundant H-bonds. 1PPT was packed by hydrophobic interactions between its long α-helix and its N-terminal and C-terminal coils. The number of turns in all D-amino acid substituted 1ppt mutants were always less than or equal to 1PPT, so the α-helix was not as easy to be expanded as β-sheet. Readjustments of dominant α-helix, N-terminal loop, and C-terminal pseudo-helix of 1ppt D-mutants were driven by new positions of hydrophobic interactions (Figure 3Biv, Bvi, and Bx). All D-amino acid monosubstituted 1ppt peptides formed equal or more H-bonds compared to 1PPT (Table S4, 1PPT worksheet tab), implying that prebiotic D-amino acids could effectively build secondary structure and pack peptides on occasions.

Besides, 1ZFI was another stable polypeptide since the 4 disulfide bonds of 1ZFI could stabilize the relative locations of its five anti-parallel β-sheets and one α-helix. Except for 1zfi_2dD and 1zfi_9dP (Figure 3Ciii and Cvii), the other 1zfi D-mutants were still rich in H-bonds between β-sheets and α-helix (Figure 3Cii, Civ, Cv, Cvi, Cviii, Cix, and Cx). Only 1zfi_5dT had the same amount but not entirely identical pairs of H-bonds compared to 1ZFI (Table S4, 1ZFI worksheet tab).

As for the relaxed structure of 2lze D-mutants, the short 3-10 helix and short β-turns were preserved constantly by H-bonds (Figure 3Dii-Dx). 2LZE had a relatively loose tertiary structure with a segment of linker and two arm-like coils, so almost all peptide bonds were always available to rotate appropriate angles and to rebuild H-bonds between proximal residues. Similar to 1ppt, all monosubstituted 2lze D-mutants possessed more H-bonds than 2LZE, except for 2lze_8dD which had only one less H-bond (Table S4, 2LZE worksheet tab).

3.5 Primordial polypeptides substituted with D-amino acid clusters

To investigate which D-amino acid clusters could stabilize folding energies or preserve original structures, each L-amino acid cluster in 1ARK, 1PPT, and 1ZFI was substituted for its counterpart cluster of D-configuration. Results demonstrated that D-amino acid mutations of the E cluster lowered 1ark's stability energy by 11.4 kcal/mol per residue; D-mutations of the LS cluster lowered 1ark's energy level by 7.74 kcal/mol per residue; the combination of E and LS clusters lowered the energy level of 1ark by 2.10 kcal/mol per residue; while all other clusters or combinations of clusters raised the stability energy of 1ark. No D-mutation of any combinations of amino acid clusters in 1ppt could decrease its stability energies. D-mutations of the D cluster decreased the stability energy of 1zfi by 2.34 kcal/mol per residue; the LS cluster stabilized 1zfi by 1.920 kcal/mol per residue; and the DE cluster combination stabilized 1zfi by 1.335 kcal/mol per residue (Table 4). In summary, acidic amino acids (i.e. D and E clusters) and LS cluster were able to decrease the folding energies of primordial polypeptides.

Besides the stability energy of peptide folding, final folded conformations also reflected the stability to preserve the original structures of proteins. Given the RMSD values shown in Table 4 that quantified the superimposition with original folded conformations, only the E cluster 1ark mutant was similar to the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

original 1ARK conformation; the 1ppt mutants of the D cluster, E cluster, IVT cluster, the combination of D and IVT clusters, the combination of E and IVT clusters, and the combination D, E, and IVT clusters were similar to the original 1PPT conformation; and the 1zfi Mutants of the D cluster, E cluster, LS cluster, IVT cluster, the combination of D and E clusters, the combination of D and LS clusters, the combination of D and IVT clusters, the combination of E and LS clusters, the combination of E and IVT clusters, the combination of D, E, and LS clusters, and the combination of D, LS, and IVT clusters were similar to the original 1ZFI conformation (Table 4).

Hitherto, the stability energy analyses determined how thermodynamically favorable the folded structure was, and RMSD values quantified how superimposable a D-mutant was with its parent all L-amino acid polypeptide. Now, we applied secondary structure occupancy analyses to understand reservable, extensible, or producible functional motifs (Table 5). The 1ark D-mutants of D, E and IVT single clusters, APDE and APEIVT three cluster combinations, and APDEIVT and APELSIVT four cluster combinations contained noticeable occupancies of β-sheet and β-turn. Regardless of what mutations it was subjected to, all 1ppt D-mutants had considerable amounts of original secondary structures, including α-helical structure and β-turn; considering the high similarity among these mutants, means ± standard errors of single or multiple cluster D-substituted 1ppt were enumerated in Table 5. The 1zfi D-mutants of AP, D, E and IVT single clusters, APLS, DIVT, ELS, EIVT and LSIVT two cluster combinations, APDLS, APELS, APEIVT, DLSIVT and ELSIVT three cluster combinations, and APDELS and APDEIVT four cluster combinations were occupied remarkably by β-sheet and helical structure (Table 5). However, the AP cluster substitution could usually remarkably alter the backbone structure of peptides, and thus it was not considered a functional domain stabilizer. In summary, D-enantiomers of D, E, and IVT were potential substitutions of their respective L-enantiomers to sustain functional domains on the basis of secondary structures.

As above, D, E, LS, and IVT clusters were reexamined in 1HX2 (characterized by short β-sheets) and 1BHI (characterized by a long α-helix and a short β-sheet) (Table 6) to test whether they could stabilize the energy or conserve the overall structure using torsion or Cartesian minimization methods. For 1hx2 mutants, almost all D, E, LS, and IVT clusters could maintain the overall structure and stability energy by torsion, except that IVT cluster increased energy slightly more than 2 kcal/mol per residue (Table 6). In addition, D-mutation of E in 1hx2 only increased minimal stability energy when keeping the original peptide structure as shown in Cartesian minimization data. For 1bhi mutants, all these 4 clusters were able to maintain the primary α-helix structure, but only D and IVT clusters maintained the short β-sheet tail attached to α-helix (Table 6). This reexamination indicated that various peptides had very different preferences to D-residues, but these clusters did optimize or conserve structure or stability to some extent.

3.6 Mutation sensitivity profiles of primordial polypeptides

All figures of mutation sensitivity profiles of 1ark, 1ppt, 1zfi, and 2lze were illustrated in Figure S2A, B, C, and D. Amongst the 1ark monosubstituted D-mutants, 9 of D-residues (26.5%) had higher ΔΔG than their counterpart L-residues, 13 of D-residues (38.2%) had lower ΔΔG than L-residues, and 12 of them (35.5%) were not changed (Figure 4A). Within the 1ppt monosubstituted D-mutants, 3 of D-residues (15.0%) had higher ΔΔG than their corresponding L-residues, 10 of D-residues (50.0%) had lower ΔΔG than L-residues, and 7 of them (35.0%) were not changed (Figure 4B). Within the 1zfi monosubstituted D-mutants, 14 of D-residues (36.8%) had higher ΔΔG than L-residues, 7 of D-residues (18.4%) had lower ΔΔG than L-residues, and 17 of them (44.7%) were unchanged (Figure 4C). For the 2lze monosubstituted D-mutants, 16 of D-residues (43.2%) had higher ΔΔG than L-residues, 10 of D-residues (27.0%) had lower

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

ΔΔG than L-residues, and 11 of them (29.7%) were unchanged (Figure 4D). According to these results, the D-amino acids that decreased ΔΔG were more abundant in 1ark and 1ppt, while the polypeptides 1zfi and 2lze that were characterized by long β-sheets, short α-helix, 3-10 helix, and long loop possessed more D-amino acids with increasing ΔΔG than those with decreasing ΔΔG. These findings indicated that the short β-sheets, long α-helix, and short loop structures showed more preferences for L-amino acids, while long β-sheets, short α-helix, 3-10 helix, and long loop showed more preferences for D-amino acids.

After combining the 4 mutation sensitivity profiles of primordial polypeptides, 42 D-residues (32.6%) were more stable than their counterpart L-residues, 40 D-residues (31.0%) were less stable than L-residues, and 47 of them (36.4%) did not alter the stability energy. Their ratios were approximately 1:1:1.

Insights of prebiotic peptide foldability and D-amino acids have also been provided by previous laboratory-based experiments. Pritsker et al. (2013) artificially synthesized HIV-1 gp41 protein composed of only D-amino acids could form the functional tertiary conformation (Pritsker et al. 1998). Tugyi et al. (2005) configured two N-terminal and three C-terminal amino acids of a mucin glycoprotein (MUC2) epitope into their counterpart D-enantiomers and found that the biological function of MUC2 was maintained (Tugyi et al. 2005). Furthermore, as inspired by a natural polymerase X from the African swine fever virus that could synthesize peptide chains by consuming racemic amino acids as the raw materials, Wang et al. (2016) successfully synthesized a D-amino acid polymerase that was functional when binding with L-DNA strands (Wang et al. 2016). Primitive peptides produced from the racemic amino acid mixtures could fold properly under the salty conditions of early Earth (Longo et al. 2013).

With regard to applications, D-amino acids enable adjustments of backbone φ and ψ angles when L-configurations are not fitting. Previous researches have incorporated D-amino acids into engineered proteins to stabilize their secondary structures, hairpins, metal-binding sites, novel conformations, and antiamyloid peptides (Hopping et al. 2014; Imperiali et al. 1992; Makwana and Mahalakshmi 2016; Peacock et al. 2009; Rana et al. 2005; Rodriguez-Granillo et al. 2011). Moreover, appropriate D-amino acid incorporations are capable of improving chemical or biological properties to biomaterial peptides, such as hydrogels (Marchesan et al. 2012). D-amino acid substituted tripeptides were experimentally determined to optimize non-covalent and intermolecular interactions in hydrogels (Garcia et al. 2018). Hence, the potential utilization of D-amino acids in life systems was further supported.

4 CONCLUSION

D-Amino acids are feasible to construct the secondary structure of peptides along with L-amino acids. Acidic amino acids (D and E), amino acids with a hydroxylic sidechain (S and T), and small hydrophobic amino acids (A and I) were possible to maintain the structures of natural proteins. The LS cluster was able to lower the folding energy of polypeptides in general; D-amino acid substitutions with single cluster D, E, or IVT are potential alternatives of their counterpart L-amino acid clusters to retain secondary structure; all the D, E, LS, and IVT clusters are potential to retain overall folded structure. Peptides with α-helices and long β-sheets, especially those stabilized by H-bonds and disulfide bonds, are usually less sensitive to D-amino acid substitutions. D-mutations were also possible to expand the length of β-sheet motifs. Polypeptides characterized by dominant loop structures, despite their variable backbone, were much more resilient than polypeptides with α-helix or β-sheet, which made them easier to preserve their functional domains. In conclusion, this computational modeling study implicates that L-amino acid homochirality may not be a prerequisite for protein structure. Namely, the origin of life is conceivable to occur prior to the origin of homochirality, and L-amino acids are not the only option for organisms. This

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

finding extends the definition of life and pinpoints the feasibility of incorporation of D-amino acids in extraterrestrial life systems.

Acknowledgement: I sincerely thank Dr. Pauline Schwartz and Dr. Carl Barratt from the University of New Haven for their guidance and supports. I also thank the University of New Haven for its support.

Conflict of Interest: Authors have no conflict of interest to declare.

Research involving Human Participants and/or Animals: I confirm that there are no human or animal participants in this research.

Informed consent: None.

References

Ashkenasy G, Jagasia R, Yadav M, Ghadiri MR (2004) Design of a directed molecular network Proc Natl Acad Sci U S A 101:10872-10877 doi:10.1073/pnas.0402674101

Bada JL, Miller SL (1987) Racemization and the Origin of Optically-Active Organic-Compounds in Living Organisms Biosystems 20:21-26 doi:10.1016/0303-2647(87)90016-5

Billard JM (2012) D-Amino acids in brain neurotransmission and synaptic plasticity Amino Acids 43:1851-1860 doi:10.1007/s00726-012-1346-3

Bourbo V, Matmor M, Shtelman E, Rubinov B, Ashkenasy N, Ashkenasy G (2011) Self-assembly and self-replication of short amphiphilic beta-sheet peptides Orig Life Evol Biosph 41:563-567 doi:10.1007/s11084-011-9257-y

Burley SK et al. (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy Nucleic Acids Res 47:D464-D474 doi:10.1093/nar/gky1004

Buschmann H, Thede R, Heller D (2000) New developments in the origins of the homochirality of biologically relevant molecules Angew Chem Int Edit 39:4033-4036 doi:10.1002/1521-3773(20001117)39:22<4033::AID-ANIE4033>3.0.CO;2-2

Canteros MG (2014) D-Arginine as a neuroprotective amino acid: promising outcomes for neurological diseases Drug Discov Today 19:627-636 doi:10.1016/j.drudis.2013.11.010

Carroll JD (2009) A New Definition of Life Chirality 21:354-358 doi:10.1002/chir.20590

Chao FA et al. (2013) Structure and dynamics of a primordial catalytic fold generated by in vitro evolution Nat Chem Biol 9:81-83 doi:10.1038/Nchembio.1138

Chaudhury S, Lyskov S, Gray JJ (2010) PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta Bioinformatics 26:689-691 doi:10.1093/bioinformatics/btq007

Daniello A, Donofrio G, Pischetola M, Daniello G, Vetere A, Petrucelli L, Fisher GH (1993) Biological Role of D-Amino-Acid Oxidase and D-Aspartate Oxidase - Effects of D-Amino Acids J Biol Chem 268:26941-26949

Fuchs SA, Berger R, Klomp LWJ, de Koning TJ (2005) D-amino acids in the central nervous system in health and disease Mol Genet Metab 85:168-180 doi:10.1016/j.ymgme.2005.03.003

Garcia AM et al. (2018) Chirality effects on peptide self-assembly unraveled from molecules to materials Chem 4:1862-1876 doi:10.1016/j.chempr.2018.05.016

Garton M, Sayadi M, Kim PM (2017) A computational approach for designing D-proteins with non-canonical amino acid optimised binding affinity Plos One 12:e0187524 doi:10.1371/journal.pone.0187524

Hammer Ø, Harper DA, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis Palaeontologia electronica 4:9

Harder E et al. (2016) OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins J Chem Theory Comput 12:281-296 doi:10.1021/acs.jctc.5b00864

Hopping G et al. (2014) Designed alpha-sheet peptides inhibit amyloid formation by targeting toxic oligomers Elife 3:e01681 doi:10.7554/eLife.01681

IBM Corp. (2019) IBM SPSS Statistics for Windows, 26.0 edn. IBM Corp., Armonk, NY

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Imperiali B, Moats RA, Fisher SL, Prins TJ (1992) A conformational study of peptides with the general structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH2: spectroscopic evidence for a peptide with significant. beta.-turn character in water and in dimethyl sulfoxide J Am Chem Soc 114:3182-3188 doi:10.1021/ja00035a002

Jackel C, Salwiczek M, Koksch B (2006) Fluorine in a native protein environment--How the spatial demand and polarity of fluoroalkyl groups affect protein folding Angew Chem Int Ed Engl 45:4198-4203 doi:10.1002/anie.200504387

Krause E, Beyermann M, Dathe M, Rothemund S, Bienert M (1995) Location of an amphipathic alpha-helix in peptides using reversed-phase HPLC retention behavior of D-amino acid analogs Anal Chem 67:252-258 doi:10.1021/ac00098a003

Krieger E, Vriend G (2014) YASARA View-molecular graphics for all devices-from smartphones to workstations Bioinformatics 30:2981-2982 doi:10.1093/bioinformatics/btu426

Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P (2015) MAESTRO - multi agent stability prediction upon point mutations Bmc Bioinformatics 16 doi:10.1186/s12859-015-0548-6

Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR (1996) A self-replicating peptide Nature 382:525-528 doi:10.1038/382525a0

Li X, Chmielewski J (2003) Peptide self-replication enhanced by a proline kink J Am Chem Soc 125:11820-11821 doi:10.1021/ja036569s

Longo LM, Blaber M (2012) Protein design at the interface of the pre-biotic and biotic worlds Arch Biochem Biophys 526:16-21 doi:10.1016/j.abb.2012.06.009

Longo LM, Lee J, Blaber M (2013) Simplified protein design biased for prebiotic amino acids yields a foldable, halophilic protein P Natl Acad Sci USA 110:2135-2139 doi:10.1073/pnas.1219530110

Makwana KM, Mahalakshmi R (2016) Capping beta-hairpin with N-terminal d-amino acid stabilizes peptide scaffold Biopolymers 106:260-266 doi:10.1002/bip.22837

Marchesan S, Easton CD, Kushkaki F, Waddington L, Hartley PG (2012) Tripeptide self-assembled hydrogels: unexpected twists of chirality Chem Commun (Camb) 48:2195-2197 doi:10.1039/c2cc16609g

Marchesan S et al. (2014) Chirality effects at each amino acid position on tripeptide self-assembly into hydrogel biomaterials Nanoscale 6:5172-5180 doi:10.1039/c3nr06752a

Miller SL (1953) A Production of Amino Acids under Possible Primitive Earth Conditions Science 117:528-529 doi:10.1126/science.117.3046.528

Miller SL (1974) Atmosphere of Primitive Earth and Prebiotic Synthesis of Amino-Acids Origins of Life and Evolution of the Biosphere 5:139-151 doi:10.1007/Bf00927019

Minervini G, Masiero A, Potenza E, Tosatto SCE (2015) Structural protein reorganization and fold emergence investigated through amino acid sequence permutations Amino Acids 47:147-152 doi:10.1007/s00726-014-1849-1

Mitchell JBO, Smith J (2003) D-amino acid residues in peptides and proteins Proteins 50:563-571 doi:10.1002/prot.10320

Morneau BN, Kubala JM, Barratt C, Schwartz PM (2014) Analysis of a chemical model system leading to chiral symmetry breaking: Implications for the evolution of homochirality J Math Chem 52:268-282 doi:10.1007/s10910-013-0261-5

Nanda V, Degrado WF (2004) Simulated evolution of emergent chiral structures in polyalanine J Am Chem Soc 126:14459-14467 doi:10.1021/ja0461825

O'Meara MJ et al. (2015) Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta J Chem Theory Comput 11:609-622 doi:10.1021/ct500864r

Paulson DJ, Shug AL (1981) Tissue Specific Depletion of L-Carnitine in Rat-Heart and Skeletal-Muscle by D-Carnitine Life Sci 28:2931-2938 doi:10.1016/0024-3205(81)90269-1

Peacock AF, Stuckey JA, Pecoraro VL (2009) Switching the chirality of the metal environment alters the coordination mode in designed peptides Angew Chem Int Ed Engl 48:7371-7374 doi:10.1002/anie.200902166

Potapov V, Cohen M, Schreiber G (2009) Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details Protein Eng Des Sel 22:553-560 doi:10.1093/protein/gzp030

Pritchett D et al. (2015) D-amino acid oxidase knockout (Dao(-/-)) mice show enhanced short-term memory performance and heightened anxiety, but no sleep or circadian rhythm disruption Eur J Neurosci 41:1167-1179 doi:10.1111/ejn.12880

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Pritsker M, Jones P, Blumenthal R, Shai YC (1998) A synthetic all D-amino acid peptide corresponding to the N-terminal sequence of HIV-1 gp41 recognizes the wild-type fusion peptide in the membrane and inhibits HIV-1 envelope glycoprotein-mediated cell fusion P Natl Acad Sci USA 95:7287-7292 doi:10.1073/pnas.95.13.7287

Pross A (2012) What is Life? What is Life?: How Chemistry Becomes Biology. Oxford University Press, Oxford

The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Rana S, Kundu B, Durani S (2005) A small peptide stereochemically customized as a globular fold with a molecular cleft Chem Commun (Camb):207-209 doi:10.1039/b413802c

Renfrew PD, Choi EJ, Bonneau R, Kuhlman B (2012) Incorporation of noncanonical amino acids into Rosetta and use in computational protein-peptide interface design Plos One 7:e32637 doi:10.1371/journal.pone.0032637

Ring D, Wolman Y, Miller SL, Friedmann N (1972) Prebiotic Synthesis of Hydrophobic and Protein Amino-Acids P Natl Acad Sci USA 69:765-768 doi:10.1073/pnas.69.3.765

Rodriguez-Granillo A, Annavarapu S, Zhang L, Koder RL, Nanda V (2011) Computational design of thermostabilizing D-amino acid substitutions J Am Chem Soc 133:18750-18759 doi:10.1021/ja205609c

Schenkelberg CD, Bystroff C (2015) InteractiveROSETTA: a graphical user interface for the PyRosetta protein modeling suite Bioinformatics 31:4023-4025 doi:10.1093/bioinformatics/btv492

Schrödinger Release 2019-4: Maestro, Schrödinger, LLC, New York, NY, 2019.

Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) The FoldX web server: an online force field Nucleic Acids Res 33:W382-W388 doi:10.1093/nar/gki387

Serrano L, Fersht AR (1989) Capping and alpha-helix stability Nature 342:296-299 doi:10.1038/342296a0

Shivakumar D, Harder E, Damm W, Friesner RA, Sherman W (2012) Improving the Prediction of Absolute Solvation Free Energies Using the Next Generation OPLS Force Field J Chem Theory Comput 8:2553-2558 doi:10.1021/ct300203w

Skolnick J, Zhou H, Gao M (2019) On the possible origin of protein homochirality, structure, and biochemical function Proc Natl Acad Sci U S A 116:26571-26579 doi:10.1073/pnas.1908241116

Soda K, Osumi T (1969) Crystalline Amino Acid Racemase with Low Substrate Specificity Biochem Bioph Res Co 35:363-368 doi:10.1016/0006-291x(69)90507-5

Storer JW, Giesen DJ, Cramer CJ, Truhlar DG (1995) Class IV charge models: a new semiempirical approach in quantum chemistry J Comput Aided Mol Des 9:87-110 doi:10.1007/bf00117280

Thiltgen G, Goldstein RA (2012) Assessing Predictors of Changes in Protein Stability upon Mutation Using Self-Consistency Plos One 7:e46084 doi:10.1371/journal.pone.0046084

Tugyi R, Uray K, Ivan D, Fellinger E, Perkins A, Hudecz F (2005) Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide P Natl Acad Sci USA 102:413-418 doi:10.1073/pnas.0407677102

Wang ZM, Xu WL, Liu L, Zhu TF (2016) A synthetic molecular system capable of mirror-image genetic replication and transcription Nat Chem 8:698-704 doi:10.1038/Nchem.2517

Wolosker H (2011) Serine racemase and the serine shuttle between neurons and astrocytes Bba-Proteins Proteom 1814:1558-1566 doi:10.1016/j.bbapap.2011.01.001

Wu M, Walker SI, Higgs PG (2012) Autocatalytic Replication and Homochirality in Biopolymers: Is Homochirality a Requirement of Life or a Result of It? Astrobiology 12:818-829 doi:10.1089/ast.2012.0819

Yang J, Spek EJ, Gong Y, Zhou H, Kallenbach NR (1997) The role of context on alpha-helix stabilization: host-guest analysis in a mixed background peptide model Protein Sci 6:1264-1272 doi:10.1002/pro.5560060614

Yun RH, Anderson A, Hermans J (1991) Proline in alpha-helix: stability and conformation studied by dynamics simulation Proteins 10:219-228 doi:10.1002/prot.340100306

Zaia DAM, Zaia CTBV, De Santana H (2008) Which Amino Acids Should Be Used in Prebiotic Chemistry Studies? Origins Life Evol B 38:469-488 doi:10.1007/s11084-008-9150-5

Zerze GH, Khan MN, Stillinger FH, Debenedetti PG (2018) Computational Investigation of the Effect of Backbone Chiral Inversions on Polypeptide Structure J Phys Chem B 122:6357-6363 doi:10.1021/acs.jpcb.8b03157

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

FIGURES

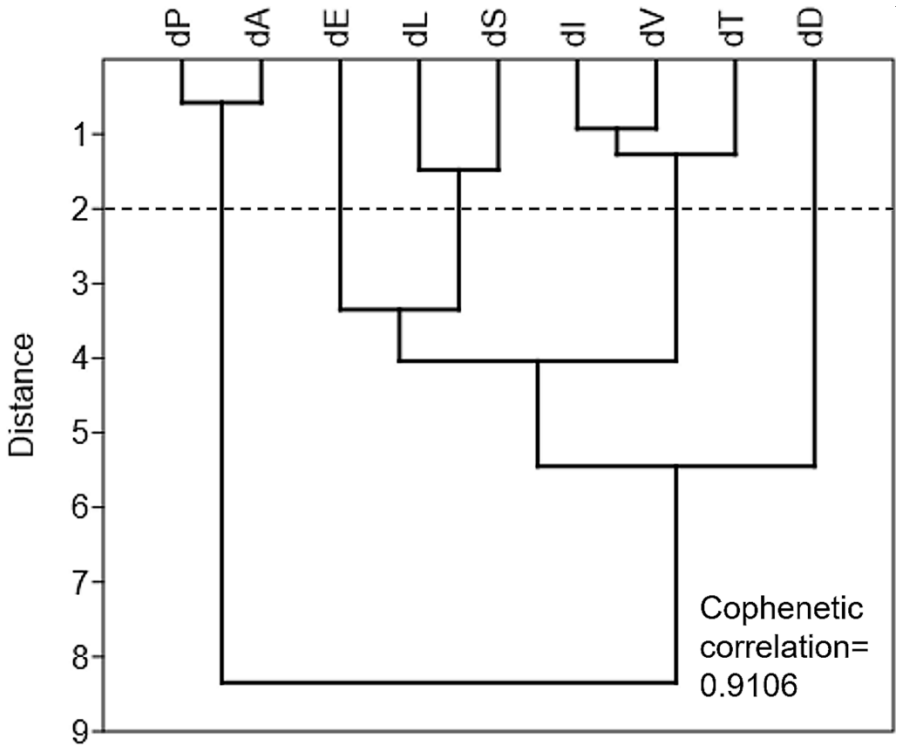


Figure 1. Hierarchical clustering (using unweighted pair group method with arithmetic mean (UPGMA) algorithm and Euclidean distance matrix) dendrogram of 9 chiral prebiotic amino acids within dipeptides.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

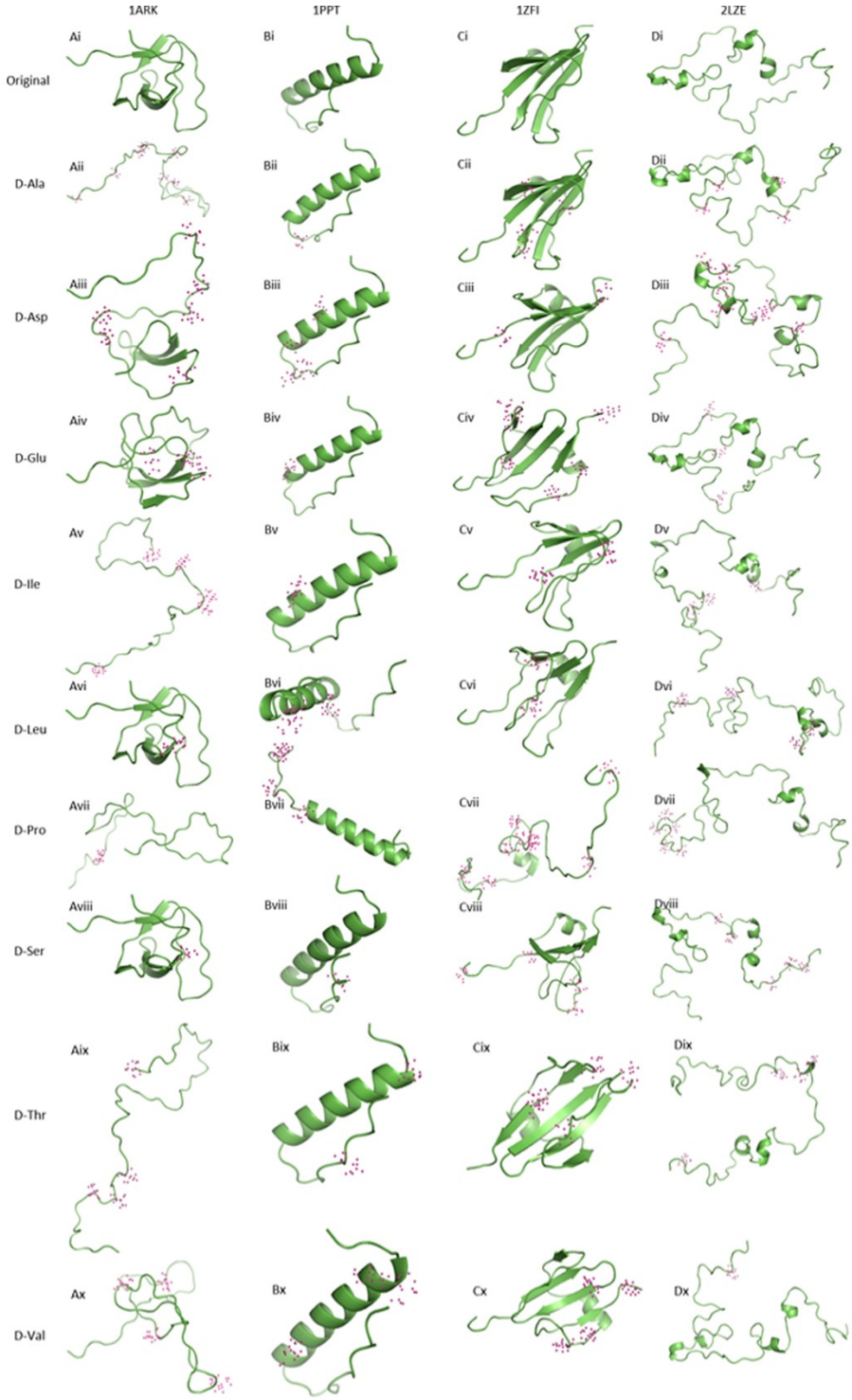


Figure 2. 1ARK (Ai-Ax), 1PPT (Bi-Bx), 1ZFI (Ci-Cx), 2LZE (Di-Dx), and their D-amino acid substituted (D-mutated) derivatives after torsion energy minimization (green ribbon, backbone chain with secondary structure depicted; pink dots, atoms in D-mutated residues). Figures A-Di, original polypeptide without D-mutations; Figures A-Dii, D-Ala mutants; Figures A-Diii, D-Asp mutants; Figures A-Div, D-Glu mutants; Figures A-Dv, D-Ile mutants; Figures A-Dvi, D-Leu mutants; Figures A-Dvii, D-Pro mutants; Figures A-Dviii, D-Ser mutants; Figures A-Dix, D-Thr mutants; Figures A-Dx, D-Val mutants.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

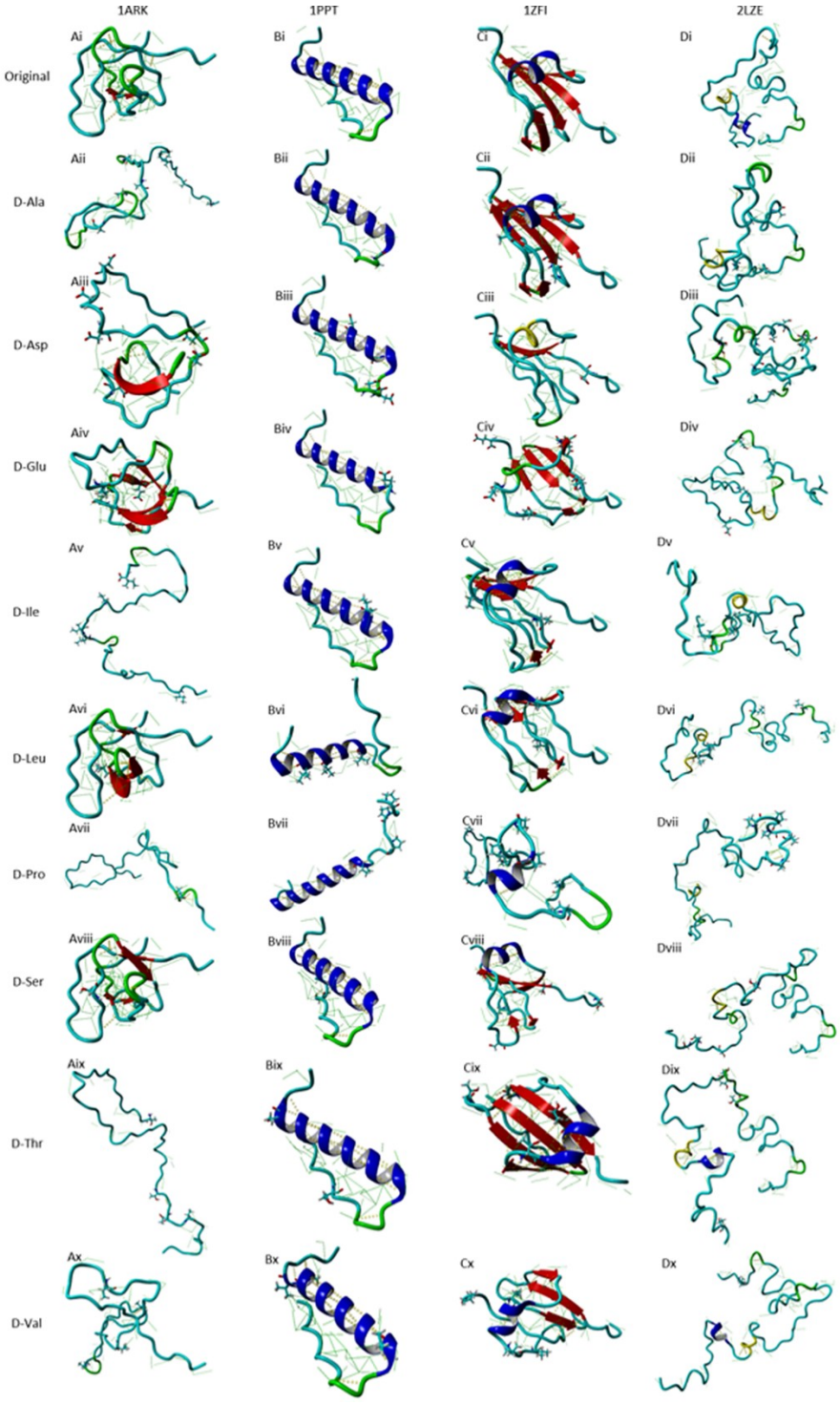


Figure 3. Hydrogen bonds and hydrophobic interactions of 1ARK (Ai-Ax), 1PPT (Bi-Bx), 1ZFI (Ci-Cx), 2LZE (Di-Dx), and their partially D-amino acid substituted derivatives after torsion energy minimization (illustrating D-mutated residues in stick style, hydrogen bonds as cylindrical yellow straight dash line, and hydrophobic interactions as green thin straight dash line). Figures A-Di, original polypeptide without D-mutations; Figures A-Dii, D-Ala mutants; Figures A-Diii, D-Asp mutants; Figures A-Div, D-Glu mutants; Figures A-Dv, D-Ile mutants; A-Dvi, D-Leu mutants; Figures A-Dvii, D-Pro mutants; Figures A-Dviii, D-Ser mutants; Figures A-Dix, D-Thr mutants; Figures A-Dx, D-Val mutants.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

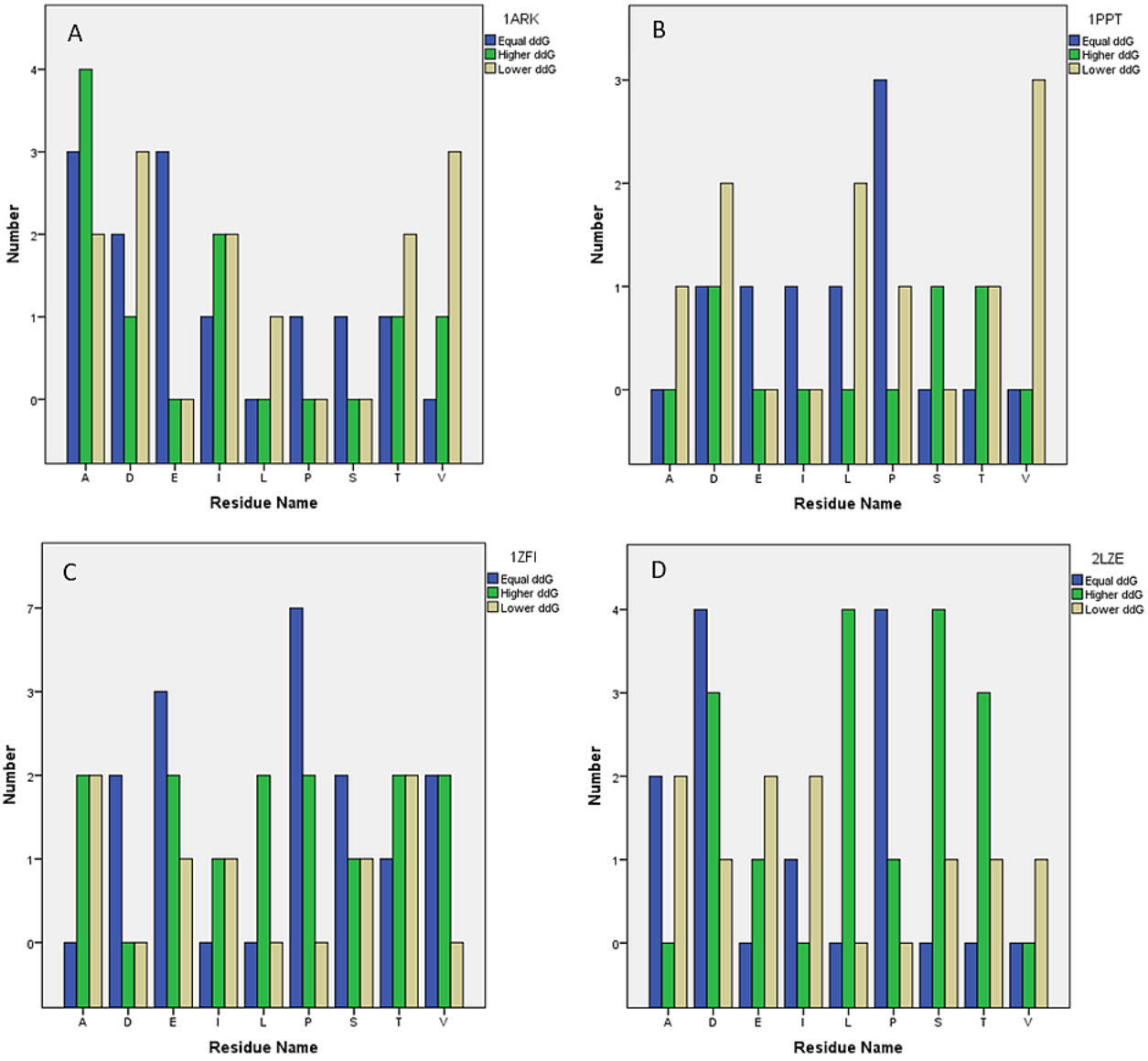


Figure 4. Numbers of D-residues with higher, lower, or equal stability energy (ΔΔG) compared to original L-residues in 1ARK (A), 1PPT (B), 1ZFI (C), and 2LZE (D) polypeptides.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

TABLES

Table 1. Sequences of 4 primordial polypeptides in computational modeling.

Polypeptide	Length (residues)	Sequence
1ARK	60	TAGKIFRAMYDYMAADADEVSFKDGDAIINVQAIDEGWMY GTVQRTGRTGMLPANYVEAI
1PPT	36	GPSQPTYPGDDAPVEDLIRFYDNLQQYLNVVTRHRY
1ZFI	67	GSHTPDESFLCYQPDQVCCFICRGAAPLPSEGECNPHPTAPW CREGAVEWVPYSTGQCRTTCIPYVE
2LZE	87	MGAPVPYPDPLEPRGGKHICAICGNNAEDYKHTDMDLTYTD RDYKNCESYHKCSDLCQYCRYQKDLAIHHQHHHGGSMGMS GSGTGY

Table 2. Characteristic secondary structure compositions in 1ARK, 1PPT, 1ZFI, and 2ZLE.

Secondary structure	1ARK	1PPT	1ZFI	2LZE
α-Helix (%)	0.0	55.6	9.4	6.9
β-Sheet (%)	21.3	0.0	44.3	0.0
β-Turn (%)	15.5	11.1	6.0	4.6
Coil or loop (%)	63.2	33.3	40.3	82.8
3-10 Helix (%)	0.0	0.0	0.0	5.7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Table 3. Effects of D-mutations of each species of prebiotic amino acid on normalized stability difference determined using the Cartesian energy minimization method.

Amino acid D-mutated:		Null	A	D	E	I	L	P	S	T	V
1ARK	Number of mutations	0	9	6	3	5	1	1	1	4	4
	Stability (kcal/mol)	41.29	85.26	74.52	46.27	54.92	63.25	77.17	38.81	99.74	60.99
	Normalized difference (kcal/mol)	-	4.89	5.54	1.66*	2.73	21.96	35.88	-2.48***	14.61	4.92
1PPT	number of mutations	0	1	4	1	1	3	4	1	2	3
	Stability (kcal/mol)	10.65	10.98	25.65	16.47	11.32	22.76	41.08	10.67	18.23	25.66
	Normalized difference (kcal/mol)	-	0.33**	3.75	5.82	0.67**	4.04	7.61	0.02**	3.79	5.00
1ZFI	Number of mutations	0	4	2	6	2	2	9	4	5	4
	Stability (kcal/mol)	70.43	96.00	86.81	75.39	95.05	95.21	138.39	74.30	70.85	96.52
	Normalized difference (kcal/mol)	-	6.39	8.19	0.83**	12.31	12.39	7.55	0.97**	0.08**	6.52
2LZE	number of mutations	0	4	8	3	3	4	5	5	4	1
	Stability (kcal/mol)	115.15	128.74	126.86	123.96	136.24	132.09	156.29	131.50	131.10	123.54
	Normalized difference (kcal/mol)	-	3.40	1.46*	2.94	7.03	4.24	8.23	3.27	3.98	8.39

*, D-residue mutation increases the stability of whole polypeptide by more than 1 kilocalorie/mole (kcal/mol) but less than 2 kcal/mol.
**, D-residue mutation increases the stability of whole polypeptide by less than 1 kcal/mol.
***, D-residue mutation decreases the stability of whole polypeptide.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Table 4. Effects of D-mutations to residue clusters on normalized stability difference and RMSD using the torsion energy minimization method compared with original all L-amino acid polypeptides (primordial polypeptides focused here were 1ARK, 1PPT, and 1ZFI).

	1ARK				1PPT				1ZFI			
	Number of mutations	Stability (kcal/mol)	Normalized difference (kcal/mol)	RMSD	Number of mutations	Stability (kcal/mol)	Normalized difference (kcal/mol)	RMSD	Number of mutations	Stability (kcal/mol)	Normalized difference (kcal/mol)	RMSD
One cluster												
AP	10	371.22	26.21	15.101	5	39.72	4.19	10.025	13	134.44	0.54	20.375
D	6	125.22	2.68	8.846	4	33.57	3.70	1.017#	2	122.73	-2.34*	0.910#
E	3	74.96	-11.39*	2.681#	1	20.60	1.83	0.957#	6	135.91	1.42	1.246#
LS	2	93.63	-7.75*	14.480	4	43.65	6.22	7.513	6	115.89	-1.92*	1.690#
IVT	13	524.79	31.97	14.837	6	88.80	11.67	3.194#	11	176.66	4.48	3.370#
Two clusters												
APD	16	120.05	0.68	24.257	9	60.38	4.62	7.773	15	141.46	0.94	25.178
APE	13	299.10	14.61	18.355	6	43.85	4.18	11.320	19	153.34	1.36	29.496
APLS	12	163.26	4.51	24.870	9	59.78	4.56	7.780	19	355.63	12.01	17.345
APIVT	23	187.88	3.42	24.149	11	72.38	4.87	13.847	24	203.60	3.17	22.137
DE	9	114.18	0.56	14.921	5	142.53	24.75	5.451	8	116.73	-1.34*	2.090#
DLS	8	111.98	0.36	22.420	8	63.54	5.60	11.571	8	140.49	1.64	1.859#
DIVT	19	320.49	11.12	18.752	10	70.92	5.22	2.883#	13	269.4	10.92	3.505#
ELS	5	98.63	-2.10*	26.552	5	58.17	7.88	14.905	12	137.41	0.83	2.444#
EIVT	16	159.09	3.12	30.908	7	51.10	4.62	2.559#	17	186.69	3.49	3.374#
LSIVT	15	204.61	6.37	22.236	10	74.59	5.58	8.249	17	162.57	2.07	5.191
Three clusters												
APDE	19	240.95	6.94	17.101	10	65.55	4.68	7.931	21	162.86	1.69	26.446
APDLS	18	130.28	1.18	21.264	13	76.12	4.41	8.660	21	329.28	9.61	16.314
APDIVT	29	194.00	2.93	25.841	15	88.00	4.62	11.570	26	347.84	8.48	15.674
APELS	15	113.76	0.31	19.223	10	65.07	4.63	9.995	25	177.89	2.02	20.125
APEIVT	26	419.38	11.93	23.623	12	81.55	5.23	11.892	30	223.71	3.21	19.645
APLSIVT	25	179.93	2.83	25.036	15	89.06	4.69	11.159	30	580.99	15.12	21.519
DELS	11	126.42	1.57	16.547	9	67.08	5.37	11.889	14	151.21	1.70	2.679#
DEIVT	22	183.09	3.36	26.106	11	87.71	6.27	3.895#	19	652.09	27.61	8.261
DLSIVT	18	179.53	3.91	19.390	15	91.90	4.88	10.565	19	225.48	5.16	2.760#
ELSIVT	21	247.63	6.60	27.226	11	78.51	5.43	8.314	23	281.23	6.69	6.464
Four clusters												
APDELS	21	319.11	10.00	15.387	14	77.82	4.22	9.447	27	205.02	2.87	17.386
APDEIVT	32	534.60	13.3	16.043	16	93.97	4.70	11.351	32	217.19	2.81	20.377
APDLSIVT	31	196.00	2.80	30.851	19	107.69	4.68	10.467	32	1210.58	33.85	16.314
APELSIVT	28	507.64	14.23	19.691	16	100.38	5.10	11.363	36	903.36	21.55	19.115
DELSIVT	24	216.92	4.49	26.344	15	101.43	5.51	9.813	25	591.52	18.56	9.410
Five clusters												
APDELSIVT	34	297.49	5.54	16.383	20	111.82	4.65	9.484	38	622.72	13.03	18.484

*, D-mutant decreases overall stability energy level

#, D-mutant with RMSD less than 4 when superimposing with all-L amino acid parent polypeptide

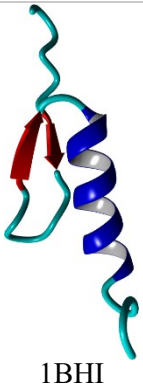
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Table 5. Secondary structure occupancies of combined residue cluster D-mutants which were analogous to their corresponding original all L-amino acid polypeptides. Numbers and amino acid single letter codes after the underscore signified the number of the following residue that was transformed from L- to D-configuration. (n cluster D-substituted 1ppt mutants were the collection of D-substituted 1ppt replaced by all combinatorial sets of n D-amino acid clusters)

	Number of mutated clusters	α-Helix (%)	β-Sheet (%)	β-Turn (%)	3-10 Helix (%)
1ARK	0	0.0	21.3	15.5	0.0
1ark_9A1P6D3E5I4V4T	4	0.0	3.3	6.7	0.0
1ark_9A1P3E1L1S5I4V4T	4	0.0	6.7	3.3	0.0
1ark_9A1P6D3E	3	0.0	3.3	6.7	0.0
1ark_9A1P3E5I4V4T	3	0.0	5.0	13.3	0.0
1ark_6D	1	0.0	18.3	18.3	0.0
1ark_3E	1	0.0	26.7	16.7	0.0
1ark_5I4V4T	1	0.0	6.7	20.0	0.0
1PPT	0	55.6	0.0	11.1	0.0
1ppt_1A4P4D1E3L1S1I3V2T	5	41.7	0.0	0.0	11.1
4 cluster D-substituted 1ppt mutants	4	43.9±3.7	0.0±0.0	6.7±5.4	5.0±6.2
3 cluster D-substituted 1ppt mutants	3	45.0±8.1	0.0±0.0	6.7±7.4	5.8±8.1
2 cluster D-substituted 1ppt mutants	2	49.7±7.6	0.0±0.0	6.7±7.4	1.4±4.2
1 cluster D-substituted 1ppt mutants	1	53.3±4.1	0.0±0.0	11.1±0.0	0.0±0.0
1ZFI	0	9.4	44.3	6.0	0.0
1zfi_4A9P2D6E2L4S	4	9.0	0.0	6.0	0.0
1zfi_4A9P2D6E2I4V5T	4	9.0	0.0	6.0	0.0
1zfi_4A9P2D2L4S	3	9.0	0.0	6.0	0.0
1zfi_4A9P6E2L4S	3	9.0	0.0	6.0	0.0
1zfi_4A9P6E2I4V5T	3	9.0	0.0	6.0	0.0
1zfi_2D2L4S2I4V5T	3	9.0	0.0	6.0	0.0
1zfi_6E2L4S2I4V5T	3	9.0	9.0	0.0	0.0
1zfi_4A9P2L4S	2	9.0	6.0	3.0	0.0
1zfi_2D2I4V5T	2	0.0	14.9	0.0	6.0
1zfi_6E2L4S	2	9.0	26.9	3.0	0.0
1zfi_6E2I4V5T	2	0.0	11.9	6.0	6.0
1zfi_2L4S2I4V5T	2	0.0	3.0	0.0	6.0
1zfi_4A9P	1	9.0	3.0	3.0	0.0
1zfi_2D	1	0.0	14.9	9.0	6.0
1zfi_6E	1	0.0	19.4	9.0	6.0
1zfi_2I4V5T	1	0.0	0.0	6.0	6.0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 January 2021

Table 6. Normalized stability difference based on Cartesian and torsion minimizations, and superimposition RMSD and identified secondary structures based on torsion minimization only for inspecting the conservative effects of suggested D-amino acid clusters.

	Peptide	Normalized difference by Cartesian minimization (kcal/mol)	Normalized difference by torsion minimization (kcal/mol)	RMSD	α-Helix (%)	β-Sheet (%)	β-Turn (%)
 1HX2	1HX2	-	-	-	0.0	18.3	33.3
	1hx2_2D	22.22	-18.32*	1.285#	0.0	16.7	20.0
	1hx2_2E	1.82*	1.28*	0.538#	0.0	16.7	33.3
	1hx2_3L7S	2.35	-0.21*	1.391#	0.0	11.7	26.7
	1hx2_1I4V5T	7.60	4.33	1.912#	0.0	0.0	13.3
 1BHI	1BHI	-	-	-	36.8	13.2	10.5
	1bhi_3D	2.44	3.07	0.413#	36.8	13.2	10.5
	1bhi_2E	5.71	6.17	11.308	34.2	0.0	10.5
	1bhi_3L1S	6.48	4.27	10.523	34.2	0.0	0.0
	1bhi_1V3T	5.02	4.81	2.318#	36.8	13.2	0.0

*, D-residue mutation increases the stability of whole polypeptide by less than 2 kcal/mol
#, D-mutant with RMSD less than 4 when superimposing with all-L amino acid parent polypeptide