

# Homochirality may not be a prerequisite for protein structure and function: a computational model investigation of prebiotic peptides

Jianxun Shen <sup>1,\*</sup>, Pauline M. Schwartz <sup>2</sup>, Carl Barratt <sup>3</sup>

<sup>1</sup> School of Earth and Environmental Sciences and Centre for Exoplanet Science, University of St Andrews, St Andrews KY16 9AL, UK

<sup>2</sup> Department of Chemistry and Chemical Engineering, Tagliatela College of Engineering, University of New Haven, West Haven, CT 06516, USA

<sup>3</sup> Department of Mechanical and Industrial Engineering, Tagliatela College of Engineering, University of New Haven, West Haven, CT 06516, USA

\* 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 homochiral and heterochiral peptides, a variety of computational methods were employed. 10 prebiotic amino acids (Gly, Ala, Asp, Glu, Ile, Leu, Pro, Ser, Thr, and Val) were previously determined by multiple previous meteorite, spark discharge, and hydrothermal vent studies. We focused on what had been reported as primary early Earth polypeptide analogs: 1ARK, 1PPT, 1ZFI, and 2LZE. Tripeptide composed of only Asp, Ser, and Val exemplified that different positions (i.e., N-terminus, C-terminus, and middle) made a difference in minimal folding energy of peptides, while the classification of amino acid (hydrophobic, acidic, or hydroxylic) did not show significant difference. Hierarchical cluster analysis for dipeptides with all possible combinations of the proposed 10 prebiotic amino acids and their D-amino acid substituted derivatives generated five clusters. Prebiotic polypeptides were built up to test the significance of molecular fluctuations, secondary structure occupancies, and folding energy differences based on these clusters. Most interestingly, among 129 residues, mutation sensitivity profiles presented that the ratio of more stable to less stable to equally stable D-amino acids was about 1:1:1. In conclusion, some combinations of a mixture of L- and D-amino acids can act as essential building blocks of life. Peptides with  $\alpha$ -helices, long  $\beta$ -sheets, and long loops are usually less sensitive to D-amino acid replacements in comparison to short  $\beta$ -sheets.

**Keywords:** D-amino acids; homochirality; prebiotic chemistry; molecular modeling

## 1 INTRODUCTION

The scientific question we intended to address and explore here was whether homochirality is the natural biosignature of life or not. 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.<sup>1</sup> The other configuration of biomolecules usually causes damages to the health of life. For instance, although L-carnitine is useful in the human body, D-carnitine interrupts the normal transportation of long-chain acyl groups in fatty acids from the mitochondrial intermembranous space into the mitochondrial matrix.<sup>2</sup> It is crucial to investigate whether homochirality is necessary for the

emergence of life on early Earth, and how primitive chemical blocks are built up for macromolecules and further for complex organisms. The evolution of biomolecules tends to minimize folding energy to stabilize their structures, a reflection of the second law of thermodynamics. However, the chemical self-replicative reactions in biological systems,<sup>3</sup> such as homochirality, are far from the favorable thermodynamic equilibrium, since thermodynamic equilibrium prefers racemic mixture.<sup>4</sup>

The origin of homochirality is not well understood and may be linked to early symmetry breaking at the beginning of the universe. According to parity violation, longitudinally polarized  $\beta$ -particles and parity-violating energy difference affect the packing energies of peptides at the nuclear and electroweak force levels.<sup>5,6</sup> The intrinsic chirality for atoms contributes to different enantiomeric energy and potentially autocatalysis.<sup>7</sup> Slight energy fluctuation could cause a “butterfly effect” during enantiomer interactions, and L-amino acids eventually won the natural selection in life systems. In addition, Carroll (2009) applied Shannon’s information entropy,  $H(S) = -\sum_{i=1}^n p_i \log_2 p_i$ , and tetrahedral conformation to explain the information capacity of biomolecules, and concluded that homochirality was a more efficient way to store and retrieve spatial information.<sup>8</sup>

In the modern world, most natural peptides with D-amino acids are usually shorter than 20 amino acids<sup>9</sup> and D-amino acids play little role in biological systems, but they sometimes are beneficial to health. Evidence shows that D-serine and D-aspartate can modulate and mitigate neurodegradation, schizophrenic symptoms, epilepsy, ischemia, and amyotrophic lateral sclerosis.<sup>10-12</sup> D-arginine protects the central nervous system from neurotoxicity and avoids the onset of mental disorders without noticeable side effects.<sup>13</sup> Although the function of D-amino acid oxidase is to prevent the harmful accumulation of D-amino acids,<sup>14</sup> knockouts of D-amino acids oxidase, which inhibits the degradation of D-amino acids, could improve spatial cognition, odor cognition, and short-term spontaneous recognition memory performance for several days.<sup>15</sup> Moreover, various amino acid racemases that transform L-amino acids to D-amino acids exist in nature.<sup>16</sup> As above, it is rational to admit that peptides with D-amino acids are potential to replace this homochiral amino acid living world.

Curiosity about the hypothesis of the first prebiotic amino acid chain expands after 1952’s Miller-Urey experiment. These early Earth spark discharge simulations consistently yielded 10 abundant racemic individual amino acids, including glycine (Gly), alanine (Ala), aspartic acid (Asp), glutamic acid (Glu), isoleucine (Ile), leucine (Leu), proline (Pro), serine (Ser), threonine (Thr), and valine (Val).<sup>17-21</sup> Among them, Gly is the only achiral amino acid, and the other 9 are chiral amino acids. A meta-analysis identically determined these 10 prebiotic amino acids based on ten previous studies of extraterrestrial comets or meteorites, hydrothermal vents, and spark discharge experiments.<sup>22</sup> All 9 chiral prebiotic amino acids have been observed in natural peptides. These natural amino acids were mostly discovered in short antibiotic or immunosuppressive peptides, such as gramicidin, actinomycin, valinomycin, tyrocidine, and cyclosporine.<sup>1,2,4,9,16</sup>

Minervini et al. (2015)<sup>23</sup> proposed ten potential prebiotic polypeptides from the Protein Data Bank (PDB) in accordance with the proteinoid theory.<sup>24,25</sup> Among them, the SH3 domain in human nebulin (PDB ID: 1ARK),<sup>26</sup> avian pancreatic polypeptide (PDB ID: 1PPT),<sup>27</sup> and leech carboxypeptidase inhibitor (PDB ID: 1ZFI)<sup>28</sup> were chosen for this study, because they cover all primary secondary structures of interest: 1ARK is a component of fibrous protein that is characterized by short  $\beta$ -sheet; 1PPT is a globular peptide characterized by long  $\alpha$ -helix; and 1ZFI is characterized by long  $\beta$ -sheet and short  $\alpha$ -helix. Additionally, to investigate the random loop structure of primordial peptides, an RNA ligase – ligase 10C (PDB ID: 2LZE) derived from *in vitro* evolution was chosen as a representative.<sup>29</sup>

In this study, we remodeled the prebiotic L-amino acids in targeted peptides into D-form and minimized the energy of new peptides. We analyzed and compared the structure and stability of each peptide and

respective “D-mutated” derivatives in terms of different aspects, including secondary structure types, packing energy, B-factor (the fluctuation of root mean square deviation, RMSD), attractive force, repulsive force, salvation energy, electrostatic potential, hydrogen bonds (H-bonds), and disulfide geometry potential. Statistical analyses were conducted to explore the mutuality between these variables. Our results implicate that some combinations of a mixture of L- and D-amino acids can act as essential building blocks for life, and heterochiral peptides are feasible in biological systems beyond Earth.

## 2 MATERIALS AND METHODS

### 2.1 Prebiotic amino acids and peptides

The 10 prebiotic amino acids focused on in this study were 6 amino acids with hydrophobic side chains (glycine, Gly, G; alanine, Ala, A; valine, Val, V; isoleucine, Ile, I; leucine, Leu, L; and proline, Pro, P), 2 acidic amino acids (aspartic acid, Asp, D; and glutamic acid, Glu, E), and 2 amino acids with the hydroxyl group on their side chain (serine, Ser, S; and threonine, Thr, T).<sup>22</sup> The 4 primordial polypeptides that cover all 10 prebiotic amino acids were 1ARK (60 residues), 1PPT (36 residues), 1ZFI (67 residues), and 2LZE (87 residues) (Table 1 and Table 2).<sup>23,29</sup> Dipeptides with all possible combinations (in total 279 dipeptides) of the 10 prebiotic L- and D-amino acids were modeled in PyMOL.<sup>30</sup> Lastly, 48 tripeptides with all combinations of Asp (the abundant prebiotic acidic amino acid), Ser (the abundant amino acid with hydroxyl group in prebiotic nature), and Val (the side chain of which represents the average effect of hydrophobic prebiotic amino acid)<sup>22</sup> were modeled.

### 2.2 Computational modeling

Maestro 11<sup>31</sup> was used for the chirality transformations of 4 polypeptides and for the energy minimization to dipeptides only. Energy minimization and root mean square deviation (RMSD) calculations were performed using InteractiveROSETTA.<sup>32</sup> The occupancies of different secondary structures were measured in YASARA (Yet Another Scientific Artificial Reality Application).<sup>33</sup> Force field of the FoldX plugin<sup>34</sup> of YASARA was used for the calculation of the stability energy ( $\Delta\Delta G$ ), the change in Gibbs free energy ( $\Delta G$ ) of the product after mutation from the starting material during peptide folding.<sup>35,36</sup> Mutation sensitive profiles of peptides were generated in the MAESTROweb (Multi AgEnt STability pRedictiOn) tool.<sup>37</sup>

### 2.3 Statistical analyses

Kruskal-Wallis independent sample tests, bivariate correlation analyses, and regression analyses were performed in IBM SPSS Statistics 25.<sup>38</sup> Hierarchical cluster analysis (using unweighted pair group method with arithmetic mean algorithm and Bray-Curtis similarity index) was performed to group 9 chiral prebiotic amino acids in PAST 3.25.<sup>39</sup>

## 3 RESULTS AND DISCUSSION

### 3.1 Secondary structure analyses

1ARK is characterized by 8 dominant anti-parallel short  $\beta$ -sheets; 1PPT is characterized by 1 dominant long  $\alpha$ -helix; 1ZFI is characterized by 5 anti-parallel  $\beta$ -sheets, 1 short  $\alpha$ -helix, and 4 disulfide bonds; and 2LZE is characterized by 2 dominant zinc loop arms and 1 short fragment of 3-10 helix (Table 3).

Herein, polypeptides substituted with one or more D-amino acids were designated using lowercase letters. For example, 1ark\_dA is the D-mutant of the all L-amino acid parent polypeptide 1ARK that switched all L-Ala residues to D-Ala.

The strategy to stabilize D-amino acids in a sequence primarily constituted by L-amino acids is to minimize the  $\chi$ -angle in peptide bonds and the placement of amino acid backbone for a minimal steric hindrance. Monte-Carlo algorithm was used in structure prediction program InteractiveROSETTA to test different least energy-driven rotational conformations to peptide chain. The 1ARK peptide was relatively unstable because only the D-mutant with three D-Glu (Figure 1Aiv), one D-Leu (Figure 1Avi), or one D-Ser (Figure 1Aviii) was capable of maintaining the original packing structure; other 6 D-mutations all deconstructed the packed conformation (Figure 1Aii, Aiii, Av, Avii, Aix, and Ax). The reason of its instability might be caused by its dominant short  $\beta$ -sheets and the absence of helical secondary structure within its original all L-configuration structure.

For 1PPT peptide, no matter which species of amino acid residue was replaced, the dominant  $\alpha$ -helix segment did not have a dramatic shape transformation, instead, only a small angular twisting at the end of helix or at the central body part (e.g., the bent  $\alpha$ -helix barrel in 1ppt\_dA, Figure 1Bii). The mutated residues not necessarily affect the shape of their proximal region, such as the D-mutations of D, I, L, T, and V (Figure 1Biii, Bv, Bvi, Bix, and Bx) within long  $\alpha$ -helical structure, while the D-mutation of A affected the original  $\alpha$ -helical structure more apparently (Figure 1Bii). Overall, the  $\alpha$ -helix was a stable secondary structure to maintain the higher levels of structure.

1ZFI has five long anti-parallel  $\beta$ -sheets and one stable helix. However, compared to 1PPT, the helical structure within 1ZFI is much shorter. Long  $\beta$ -sheets in D-amino acid substituted 1ZFI could be divided into several short sheets to stabilize the functional structure (e.g., Figure 1Cix).

2LZE is a special primordial polypeptide for its high flexibility of two long loops: 20 different structural models can be found in PDB database, and their RMSD values are significantly different from each other. All D-amino acid substituted 2LZEs shared similar loop-turn-helix-loop structure with two obvious zinc arms (Figure 1Dii-Dx), indicating the maintenance of its function in these heterochiral amino acid sequences.

In all cases of the polypeptides examined, proline made a big difference to original structure because of its rigid ring structure. Taken all together, peptides with  $\alpha$ -helices, long  $\beta$ -sheets, and long loops were less sensitive to D-amino acid replacements in comparison to those with short  $\beta$ -sheets.

In this research, the effect of D-mutations on stability was normalized by the following equation:

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

The minimization approach by optimizing the torsion energy predicts folding structures without concerning the function domains in original polypeptides. Here, we reserved the functional structure of polypeptides after D-mutations and determined the stability energy difference from their respective original all L-amino acid polypeptides. The minimization approach by optimizing the placement of original Cartesian coordinates can achieve this goal. The cartesian energy minimization is less flexible on angle constraint and torsion but predict the minimal folding energy while retaining the original tertiary structure. By Cartesian energy minimization, several D-amino acid substituted polypeptides had equivalent folding energies to their corresponding original all L-amino acid polypeptides. They are 1ark\_dE, 1ark\_dS, 1ppt\_dA, 1ppt\_dI, 1ppt\_dS, 1zfi\_dE, 1zfi\_dS, 1zfi\_dT, and 2lze\_dD (Table 4).

### 3.2 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 are essential to peptide conformation were analyzed in YASARA. Hydrophobic interactions from hydrocarbon chains universally participated in the intramolecular interactions between residues (Figure 2).

1ARK polypeptide has 8 anti-parallel  $\beta$ -sheets connected by H-bonds. The tertiary structure of 1ARK is packed by multiple H-bonds between anti-parallel  $\beta$ -sheets and multiple hydrophobic interactions throughout its primary sequence (Figure 2Ai). Polypeptides 1ark\_dD, 1ark\_dE, 1ark\_dL, and 1ark\_dS had expanded  $\beta$ -sheets compared to 1ARK (Figure 2Aiii, Aiv, Avi, and Aviii). This finding suggests that D-mutations were possible to produce the long  $\beta$ -sheet motif. However, in 1ark\_dA, 1ark\_dL, 1ark\_dP, 1ark\_dT, and 1ark\_dV, H-bonds only occurred at some of turning points of the loop structure because of their less packed tertiary structure (Figure 2Aii, Av, Avii, Aix, and Ax).

The  $\alpha$ -helix of 1PPT is stabilized by abundant H-bonds. 1PPT is packed by hydrophobic interactions between its long  $\alpha$ -helix and its N-terminal and C-terminal coils. The  $\alpha$ -helix was not as easy to be expanded as  $\beta$ -sheet, because the number of turns in all D-amino acid substituted 1ppt mutants were always less than or equal to 1PPT. Readjustments of dominant  $\alpha$ -helix, N-terminal loop, and C-terminal pseudo-helix of 1ppt were motivated by new positions of hydrophobic interactions (Figure 2Biv, Bvi, and Bx).

1ZFI is another stable polypeptide because the 4 disulfide bonds of 1ZFI can stabilize the relative locations of its five anti-parallel  $\beta$ -sheets and one  $\alpha$ -helix, and thus its tertiary structure. Except 1zfi\_dD and 1zfi\_dP (Figure 2Ciii and Cvii), other 1zfi D-mutants were still rich in H-bonds between  $\beta$ -sheets and  $\alpha$ -helix (Figure 2Cii, Civ, Cv, Cvi, Cviii, Cix, and Cx).

For the relaxed structure of 2lze D-mutants, the short 3-10 helix and short  $\beta$ -turns were preserved constantly by H-bonds (Figure 2Dii-Dx). 2LZE has relatively loose tertiary structure with a segment of linker and two arm-like coils, so it is always convenient to twist necessary peptide bonds and rebuild H-bonds between proximal residues.

### 3.3 Correlation and regression analyses between the secondary structure occupancy and stability

Secondary structure occupancies and stability energies of 4 prebiotic polypeptides and their D-mutants were determined by YASARA and FoldX (Table 5). Results from the correlation analysis showed stability energy was positively correlated with the number of mutations (Pearson correlation  $r=0.42^{***}$ ,  $p<0.001$ ) and the occupancy of coil ( $r=0.44^{***}$ ,  $p<0.001$ ), while negatively correlated with  $\beta$ -sheet occupancy ( $r=-0.29^{**}$ ,  $p=0.002$ ). Stability energy was negatively correlated with  $\alpha$ -helix ( $r=-0.014$ ,  $p=0.09$ ) and  $\beta$ -turn occupancies ( $r=-0.069$ ,  $p=0.26$ ), and positively correlated with 3-10 helix occupancy ( $r=0.074$ ,  $p=0.24$ ), although not statistically significant (Table 5).

Automatic linear modeling analysis output the linear model: Stability =  $8.7 \times$  number of mutations –  $88 \times$   $\beta$  sheet occupancy +  $135 \times$  coil occupancy + 1.2. Stability was significantly associated with number of mutations (importance=  $0.441^{***}$ ,  $p<0.001$ ), coil occupancy (importance=  $0.427^{***}$ ,  $p<0.001$ ), and  $\beta$ -sheet occupancy (importance=  $0.132^*$ ,  $p<0.024$ ).

These results indicate that  $\alpha$ -helix,  $\beta$ -sheet, and  $\beta$ -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.

### 3.4 Explorations to modeled oligopeptides

I modeled multiple tripeptides with Asp (acidic), Ser (hydroxylic), and Val (hydrophobic) residues only, including DSV, DVS, SDV, SVD, VDS, and VSD, as well as different combinations of their D-amino acid substituted derivatives. We found that different positions (i.e., N-terminus, C-terminus, or middle) of D-mutations made a difference in the folding energy of peptides (Kruskal-Wallis test statistic= 6.22\*,  $p=0.045$ ). However, the chemical property-based classification (acidic, hydroxyl, or hydrophobic) of amino acids did not show significant difference (Kruskal-Wallis test statistic= 3.59,  $p=0.17$ ) without changing positions in these tripeptides.

Next, different 99 combinations of dipeptides composed of prebiotic amino acids (Gly, Ala, Asp, Glu, Ile, Leu, Pro, Ser, Thr, and Val) and their 180 combinations of D-amino acid substituted derivatives were built up. The hierarchical cluster analysis group these 10 prebiotic amino acids into five clusters – (1) Ala and Pro, (2) Asp, (3) Glu, (4) Leu and Ser, and (5) Ile, Val, and Thr (Figure 3). In the five clusters, A and P are small hydrophobic amino acids, and the D-mutations of the AP cluster caused a decrease in energy by 0.057 kilocalorie/mole (kcal/mol); D is a small acidic amino acid, and the D-mutations of the D cluster caused an increase in energy by 1.17 kcal/mol; E is a large acidic amino acid, and the D-mutation of the E cluster caused an increase in energy by 1.42 kcal/mol; L and S are large hydrophobic or small hydroxylic amino acids, and the D-mutation of the LS cluster caused an increase in energy by 1.42 kcal/mol; and I, V, and T are large or middle hydrophobic or large hydroxylic amino acids, and the D-mutation of the IVT cluster caused an increase in energy by 2.20 kcal/mol.

### 3.5 Prebiotic polypeptides substituted with D-amino acid clusters

To investigate which D-amino acid cluster can stabilize the folding energy level or maintain the functional motif structure, we substituted each L-amino acid cluster with its counterpart cluster of D-configuration in 1ARK, 1PPT, and 1ZFI. The reason to exclude 2LZE was that 2LZE and its D-amino acid substituted mutants are too variable to determine their stability and secondary structure constitution, so the result would not be reliable.

Results demonstrated that D-amino acid mutations of the E cluster lowered 1ark's energy level 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, but all other clusters or combinations of clusters enlarged the folding energy level of 1ark.

As for 1ppt, no D-mutation of any combinations of amino acid clusters could decrease the stability energies of this polypeptide.

D-mutations of the D cluster was able to decrease the folding energy level of 1zfi by 2.34 kcal/mol per residue. The LS cluster was able to stabilize 1zfi by 1.920 kcal/mol per residue, and the DE cluster combination was able to stabilize 1zfi by 1.335 kcal/mol per residue (Table 6). This result indicated that acidic amino acids (i.e. Asp and Glu) and Leu/Ser cluster were able to lower the folding energy level of polypeptides.

Besides the folding energy-based stability, the overall folded conformation is also important, because it reflects the functional structure of proteins. The folded conformation can be quantified by RMSD. Given RMSD values as displayed in Table 6, only the E cluster mutant was similar to the original 1ARK conformation.

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 all close to the original 1PPT conformation.

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 all similar to the original 1ZFI conformation.

Hitherto, the stability energy analyses determined how thermodynamically favorable the folded structure is, and RMSD values quantified how superimposable a D-mutant is with its functional parent all L-amino acid polypeptide. Now, we applied secondary structure occupancy analyses to understand reservable, extensible, or producible functional motifs. The 1ark mutants with D-mutations of D, E and IVT single clusters, APDE and APEIVT three cluster combinations, and APDEIVT and APELSIVT four cluster combinations contained noticeable occupancies of  $\beta$ -sheet and  $\beta$ -turn.

All 1ppt D-mutants had considerable amount of functional secondary structure, including helical structure and  $\beta$ -turn, regardless of what mutations it was subjected to.

The 1zfi mutants with D-mutations 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  $\beta$ -sheet and helical structure. Overall, D-enantiomers of APDEIVT, APEIVT, D, E, and IVT were potential replacements of their respective L-enantiomers.

### 3.6 Mutation sensitivity profiles of prebiotic polypeptides

The mutation sensitivity profile illustrates the change in Gibbs' free energy ( $\Delta\Delta G$ ) of a D-amino acid after being replaced by all 20 proteinogenic L-amino acids, so the higher the  $\Delta\Delta G$  in mutation sensitivity profile is, the more stable the D-amino acid substituted polypeptide is. In the result of 1ark polypeptide, 9 (26.5%) of D-residues had higher  $\Delta\Delta G$  than their counterpart L-residues, 13 (38.2%) of D-residues had lower  $\Delta\Delta G$  than L-residues, and 12 (35.5%) of them were unchanged (Figure 4A). In the result of 1ppt polypeptide, 3 (15.0%) of D-residues had higher  $\Delta\Delta G$  than their corresponding L-residues, 10 (50.0%) of D-residues had lower  $\Delta\Delta G$  than L-residues, and 7 (35.0%) of them were unchanged (Figure 4B). D-amino acids that decrease  $\Delta\Delta G$  dominate the primary sequences of 1ark and 1ppt, the polypeptides with short  $\beta$ -sheets, long  $\alpha$ -helix, and short loop.

In the result of 1zfi polypeptide, 14 (36.8%) of D-residues had higher  $\Delta\Delta G$  than L-residues, 7 (18.4%) of D-residues had lower  $\Delta\Delta G$  than L-residues, and 17 (44.7%) of them were unchanged (Figure 4C). For 2lze polypeptide, 16 (43.2%) of D-residues had higher  $\Delta\Delta G$  than L-residues, 10 (27.0%) of D-residues had lower  $\Delta\Delta G$  than L-residues, and 11 (29.7%) of them were unchanged (Figure 4D). Polypeptides 1zfi and 2lze that were characterized by long  $\beta$ -sheets, short  $\alpha$ -helix, 3-10 helix, and long loop possessed more D-amino acids with increasing  $\Delta\Delta G$  than those with decreasing  $\Delta\Delta G$ . These findings indicate that the short  $\beta$ -sheets, long  $\alpha$ -helix, and short loop structures preferred L-amino acids, while long  $\beta$ -sheets, short  $\alpha$ -helix, 3-10 helix, and long loop preferred D-amino acids.

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

### 3.7 Synthetic D-peptides

Insights of prebiotic peptide foldability and D-amino acids are also provided by previous laboratory-based experiments. Pritsker et al. (2013) artificially synthesized HIV-1 gp41 protein composed of only D-amino acids can fold properly<sup>40</sup> under salty conditions on early Earth.<sup>41</sup> Tugyi et al. (2005) turned two N-terminal and three C-terminal amino acids of a mucin glycoprotein (MUC2) epitope into their counterpart D-enantiomers, and found that the biotic function of MUC2 was maintained.<sup>42</sup> Furthermore, as inspired by a natural polymerase X from the African swine fever virus that could synthesize peptide chain by using racemic amino acids as the raw materials, Wang et al. (2016) successfully synthesized a D-amino acid polymerase that was functional to L-DNA strands.<sup>43</sup> Hence, the potential of D-amino acids in life systems is further supported.

## 4 CONCLUSION

D-Amino acids are possible to construct secondary structure with L-amino acids. D-mutations to Glu and Ser on 1ARK, Ala, Ile, and Ser on 1PPT, Glu, Ser, and Thr on 1ZFI, and Asp on 2LZE did not change the stability markedly in comparison to their corresponding original all L-amino acid polypeptides when using Cartesian energy minimization method. This result indicated that acidic amino acids (Glu and Asp), amino acids with hydroxyl group (Ser and Thr), and small hydrophobic amino acids (Ala and Ile) were possible to maintain the same structure and function as the current natural proteins. In addition, Asp, Glu, Leu, and Ser residues could also expand short  $\beta$ -sheets in 1ark. Peptides with  $\alpha$ -helix and long  $\beta$ -sheet, especially those stabilized by abundant H-bonds within parallel or antiparallel  $\beta$ -sheet cluster, are usually less sensitive to D-amino acid replacements. D-mutations were possible to produce long  $\beta$ -sheet motif, while  $\alpha$ -helix was not as easy to be expanded as  $\beta$ -sheet. Tertiary structure could be stabilized by abundant disulfide bonds, such as the four disulfide bonds on 1ZFI. Functional polypeptides characterized by dominant loop occupancy, such as 2LZE, in spite of the less stable and more mobile backbone, were much more flexible than polypeptides with observable  $\alpha$ -helix or  $\beta$ -sheet, which made them easier to preserve their functional domains.

Some combinations of D-amino acid clusters could generate functional polypeptides with L-amino acids. Based on the analyses about stability energy level, secondary structure occupancy, and RMSD of relative position, LS cluster was able to lower the folding energy level of polypeptides in general; D-amino acid substitutions by four clusters APDEIVT, three clusters APEIVT, or single cluster D, E, or IVT were potential alternatives of their counterpart L-amino acid clusters to retain secondary structure; all of D, E, LS, and IVT clusters were potential to generally retain overall folded structure even via torsion energy minimization method.

Additionally in the tripeptide trial, it displayed that position of residues played a role in stability determination, but the classification of residues did not make remarkable difference in stability. Dipeptide modeling disclosed that in short peptides, D-alanine and D-proline were conceivable to decrease the energy level after energy minimization. This computational modeling study implicate that homochirality may not be a prerequisite for functional protein structure. Namely, the origin of life is possible to occur prior to the origin of homochirality, and L-amino acids may not be the only option for complex organism biosignatures. This finding extends the definition of life and pinpoints the possibility of heterochirality in extraterrestrial life systems.

## Acknowledgements

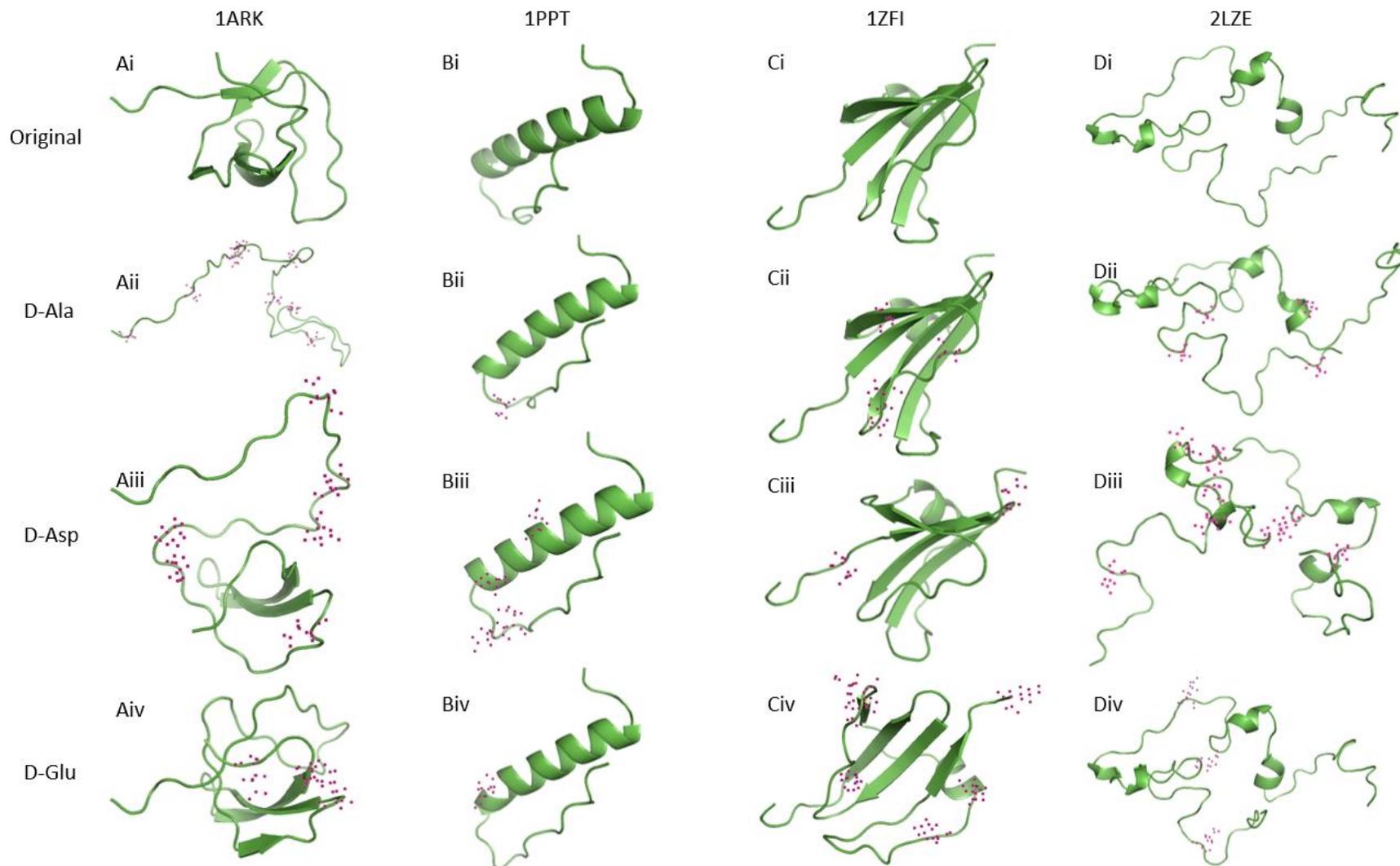
We sincerely thank Dr. Ali Senejani and Dr. Christina Zito from the Department of Biology and Environmental Science, College of Arts and Sciences, University of New Haven for their supports.

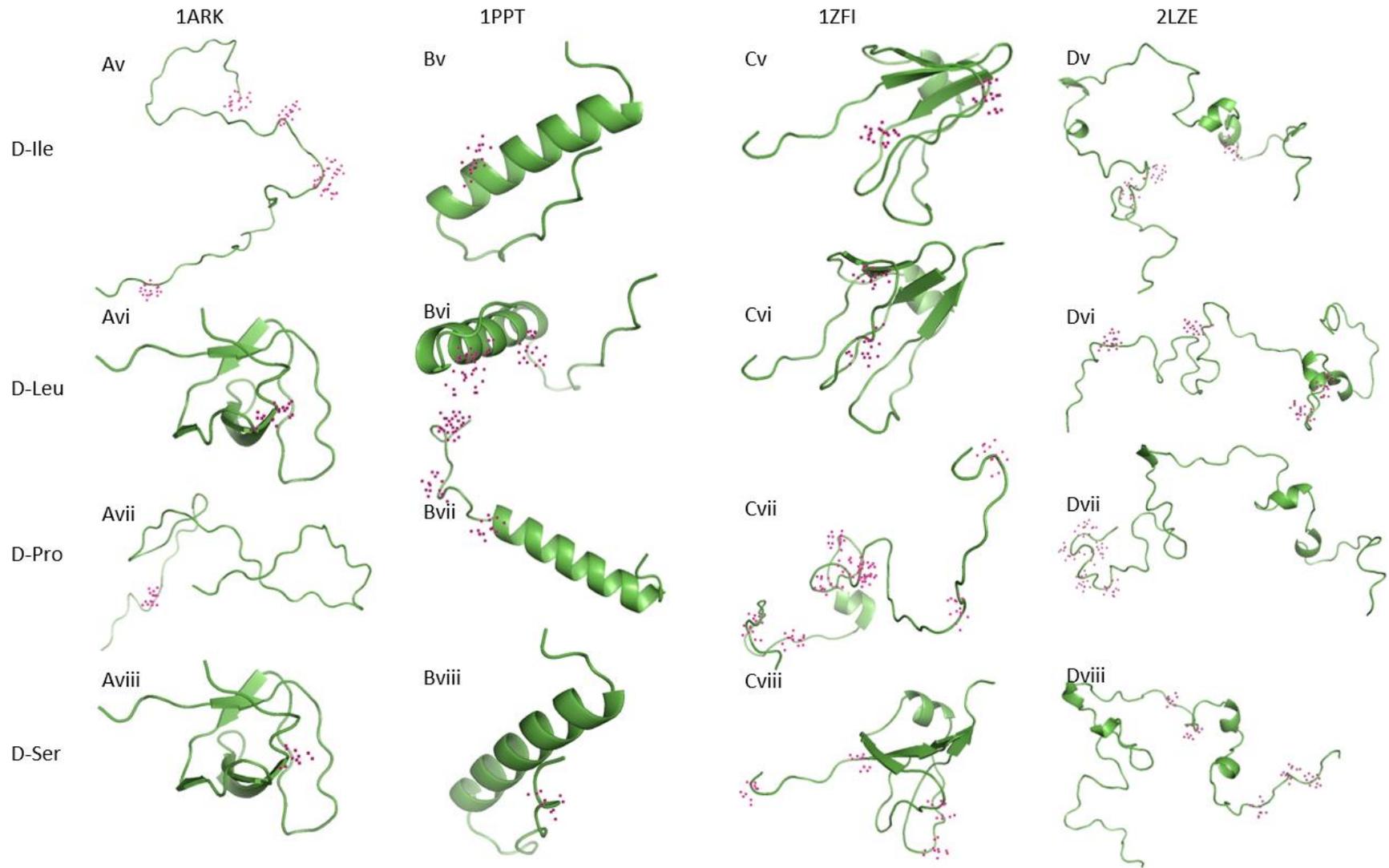
## References

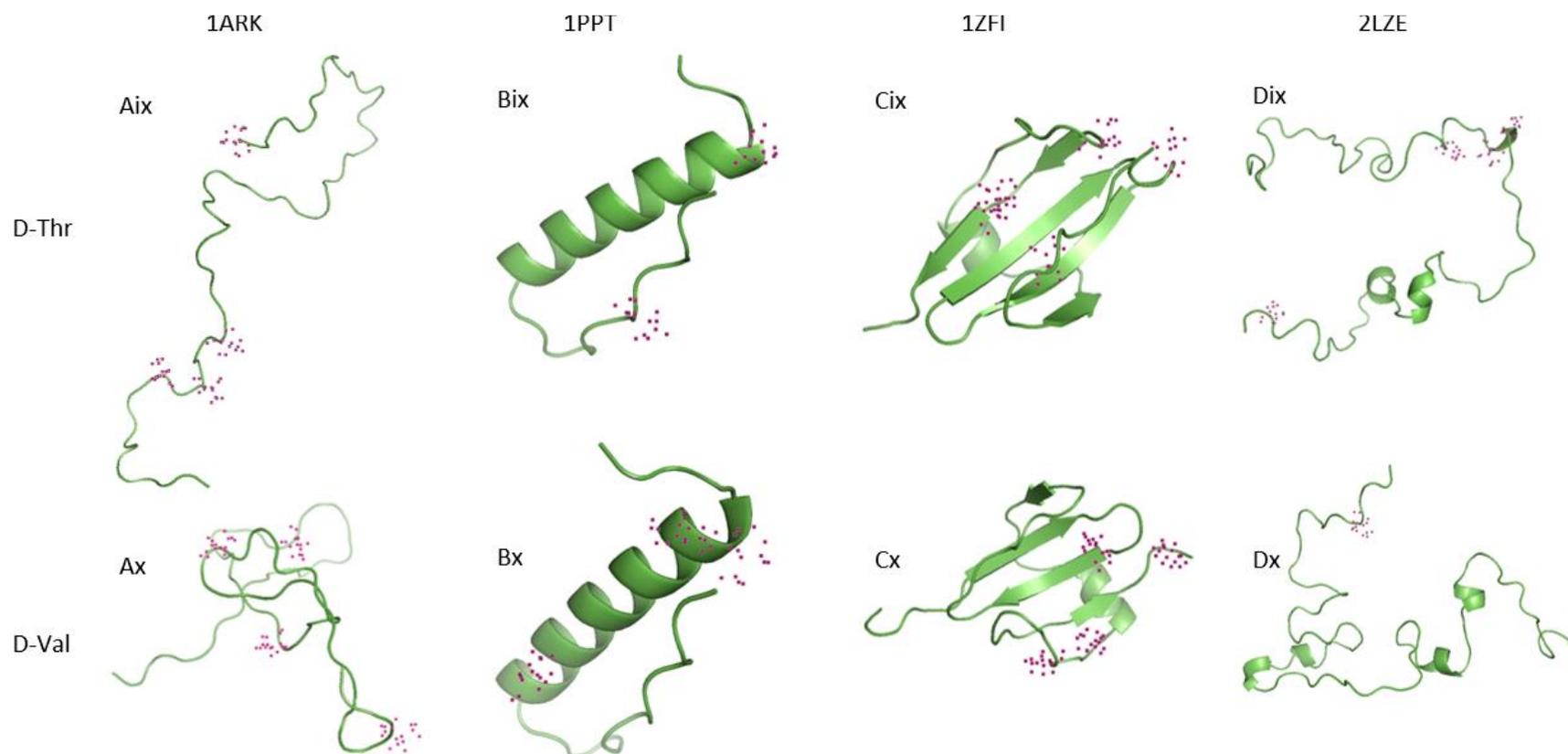
1. Morneau BN, Kubala JM, Barratt C, Schwartz PM. Analysis of a chemical model system leading to chiral symmetry breaking: Implications for the evolution of homochirality. *Journal of Mathematical Chemistry* 2014;52(1):268-282.
2. Paulson DJ, Shug AL. Tissue Specific Depletion of L-Carnitine in Rat-Heart and Skeletal-Muscle by D-Carnitine. *Life Sciences* 1981;28(26):2931-2938.
3. Wu M, Walker SI, Higgs PG. Autocatalytic Replication and Homochirality in Biopolymers: Is Homochirality a Requirement of Life or a Result of It? *Astrobiology* 2012;12(9):818-829.
4. Pross A. *What is Life?* Oxford: Oxford University Press; 2012.
5. Bonner WA. Parity violation and the evolution of biomolecular homochirality. *Chirality* 2000;12(3):114-126.
6. Poline M, Rebrov O, Larsson M, Zhaunerchyk V. Theoretical studies of infrared signatures of proton-bound amino acid dimers with homochiral and heterochiral moieties. *Chirality* 2020.
7. Buschmann H, Thede R, Heller D. New developments in the origins of the homochirality of biologically relevant molecules. *Angewandte Chemie-International Edition* 2000;39(22):4033-4036.
8. Carroll JD. A New Definition of Life. *Chirality* 2009;21(3):354-358.
9. Mitchell JBO, Smith J. D-amino acid residues in peptides and proteins. *Proteins-Structure Function and Genetics* 2003;50(4):563-571.
10. Fuchs SA, Berger R, Klomp LWJ, de Koning TJ. D-amino acids in the central nervous system in health and disease. *Molecular Genetics and Metabolism* 2005;85(3):168-180.
11. Billard JM. D-Amino acids in brain neurotransmission and synaptic plasticity. *Amino Acids* 2012;43(5):1851-1860.
12. Wolosker H. Serine racemase and the serine shuttle between neurons and astrocytes. *Biochimica Et Biophysica Acta-Proteins and Proteomics* 2011;1814(11):1558-1566.
13. Canteros MG. D-Arginine as a neuroprotective amino acid: promising outcomes for neurological diseases. *Drug Discovery Today* 2014;19(5):627-636.
14. Daniello A, Donofrio G, Pischetola M, Daniello G, Vetere A, Petrucelli L, Fisher GH. Biological Role of D-Amino Acid Oxidase and D-Aspartate Oxidase - Effects of D-Amino Acids. *Journal of Biological Chemistry* 1993;268(36):26941-26949.
15. Pritchett D, Hasan S, Tam SKE, Engle SJ, Brandon NJ, Sharp T, Foster RG, Harrison PJ, Bannerman DM, Peirson SN. D-amino acid oxidase knockout (Dao<sup>-/-</sup>) mice show enhanced short-term memory performance and heightened anxiety, but no sleep or circadian rhythm disruption. *European Journal of Neuroscience* 2015;41(9):1167-1179.
16. Soda K, Osumi T. Crystalline Amino Acid Racemase with Low Substrate Specificity. *Biochemical and Biophysical Research Communications* 1969;35(3):363-&.
17. Bada JL, Miller SL. Racemization and the Origin of Optically-Active Organic-Compounds in Living Organisms. *Biosystems* 1987;20(1):21-26.
18. Miller SL. A Production of Amino Acids under Possible Primitive Earth Conditions. *Science* 1953;117(3046):528-529.
19. Miller SL. Atmosphere of Primitive Earth and Prebiotic Synthesis of Amino-Acids. *Origins of Life and Evolution of the Biosphere* 1974;5(1-2):139-151.
20. Zaia DAM, Zaia CTBV, De Santana H. Which Amino Acids Should Be Used in Prebiotic Chemistry Studies? *Origins of Life and Evolution of Biospheres* 2008;38(6):469-488.
21. Ring D, Wolman Y, Miller SL, Friedmann N. Prebiotic Synthesis of Hydrophobic and Protein Amino-Acids. *Proceedings of the National Academy of Sciences of the United States of America* 1972;69(3):765-+.
22. Longo LM, Blaber M. Protein design at the interface of the pre-biotic and biotic worlds. *Archives of Biochemistry and Biophysics* 2012;526(1):16-21.
23. Minervini G, Masiero A, Potenza E, Tosatto SCE. Structural protein reorganization and fold emergence investigated through amino acid sequence permutations. *Amino Acids* 2015;47(1):147-152.
24. Turian G. Origin of life. II. From prebiotic replicators to protocells. *Archives Des Sciences* 1999;52(2):101-109.

25. Burley SK, Berman HM, Bhikadiya C, Bi CX, Chen L, Di Costanzo L, Christie C, Dalenberg K, Duarte JM, Dutta S and others. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Research* 2019;47(D1):D464-D474.
26. Politou AS, Millevoi S, Gautel M, Kolmerer B, Pastore A. SH3 in muscles: Solution structure of the SH3 domain from nebulin. *Journal of Molecular Biology* 1998;276(1):189-202.
27. Blundell TL, Pitts JE, Tickle IJ, Wood SP, Wu CW. X-Ray-Analysis (1.4-Å Resolution) of Avian Pancreatic-Polypeptide - Small Globular Protein Hormone. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* 1981;78(7):4175-4179.
28. Arolas JL, D'Silva L, Popowicz GM, Aviles FX, Holak TA, Ventura S. NMR structural characterization and computational predictions of the major intermediate in oxidative folding of leech carboxypeptidase inhibitor. *Structure* 2005;13(8):1193-1202.
29. Chao FA, Morelli A, Haugner JC, Churchfield L, Hagmann LN, Shi L, Masterson LR, Sarangi R, Veglia G, Seelig B. Structure and dynamics of a primordial catalytic fold generated by in vitro evolution. *Nature Chemical Biology* 2013;9(2):81-83.
30. The PyMOL Molecular Graphics System VS, LLC.
31. Schrödinger Release 2019-4: Maestro S, LLC, New York, NY, 2019.
32. Schenkelberg CD, Bystroff C. InteractiveROSETTA: a graphical user interface for the PyRosetta protein modeling suite. *Bioinformatics* 2015;31(24):4023-4025.
33. Krieger E, Vriend G. YASARA View-molecular graphics for all devices-from smartphones to workstations. *Bioinformatics* 2014;30(20):2981-2982.
34. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic Acids Research* 2005;33:W382-W388.
35. Thiltgen G, Goldstein RA. Assessing Predictors of Changes in Protein Stability upon Mutation Using Self-Consistency. *Plos One* 2012;7(10).
36. Potapov V, Cohen M, Schreiber G. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Engineering Design & Selection* 2009;22(9):553-560.
37. Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P. MAESTRO - multi agent stability prediction upon point mutations. *Bmc Bioinformatics* 2015;16.
38. IBM SPSS Statistics for Windows. 25.0. Armonk, NY: IBM Corp.; 2017.
39. Hammer Ø, Harper DA, Ryan PD. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia electronica* 2001;4(1):9.
40. Pritsker M, Jones P, Blumenthal R, Shai YC. 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. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95(13):7287-7292.
41. Longo LM, Lee J, Blaber M. Simplified protein design biased for prebiotic amino acids yields a foldable, halophilic protein. *Proceedings of the National Academy of Sciences of the United States of America* 2013;110(6):2135-2139.
42. Tugyi R, Uray K, Ivan D, Fellingner E, Perkins A, Hudecz F. Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(2):413-418.
43. Wang ZM, Xu WL, Liu L, Zhu TF. A synthetic molecular system capable of mirror-image genetic replication and transcription. *Nature Chemistry* 2016;8(7):698-704.

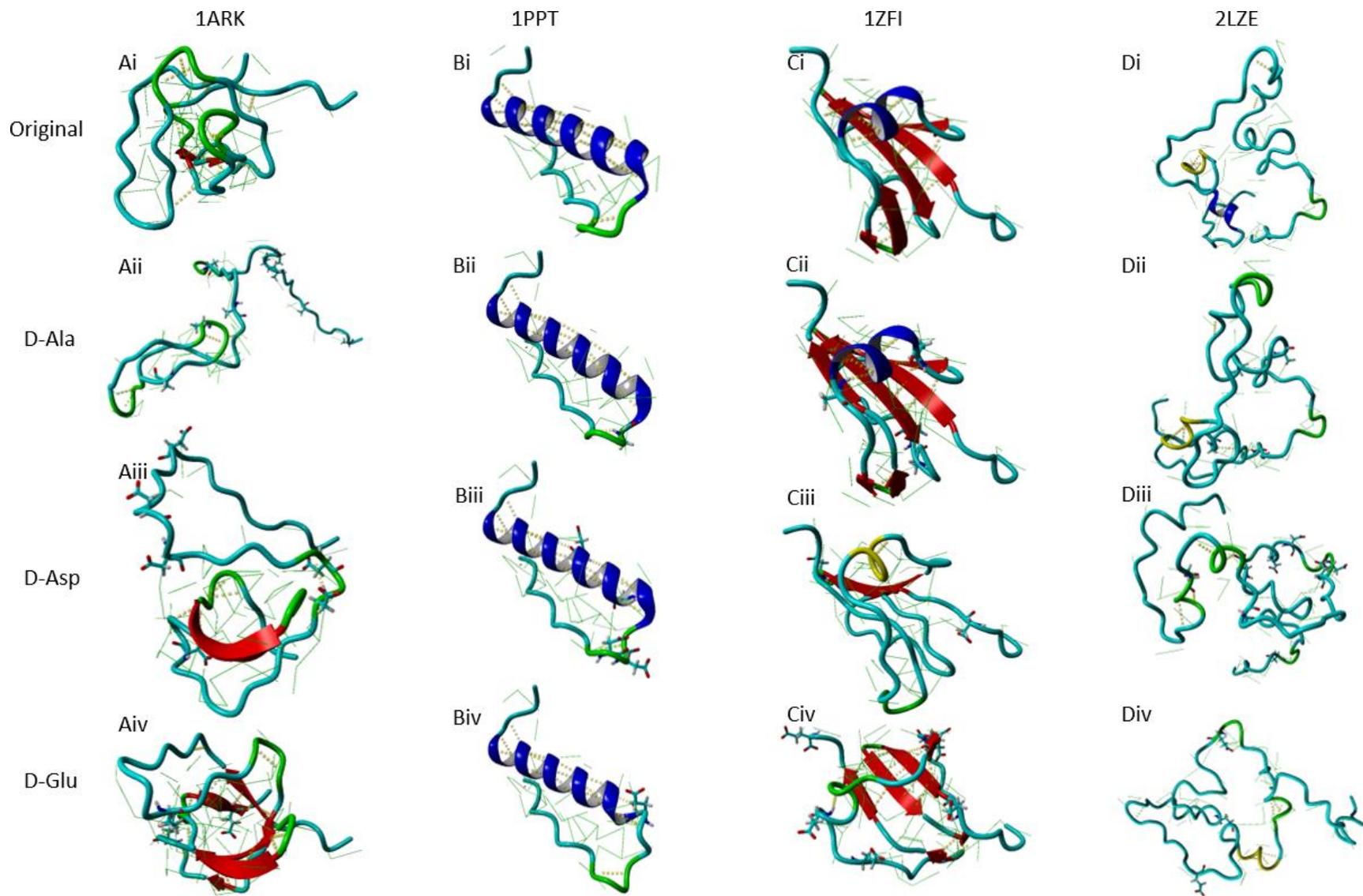
## Figures

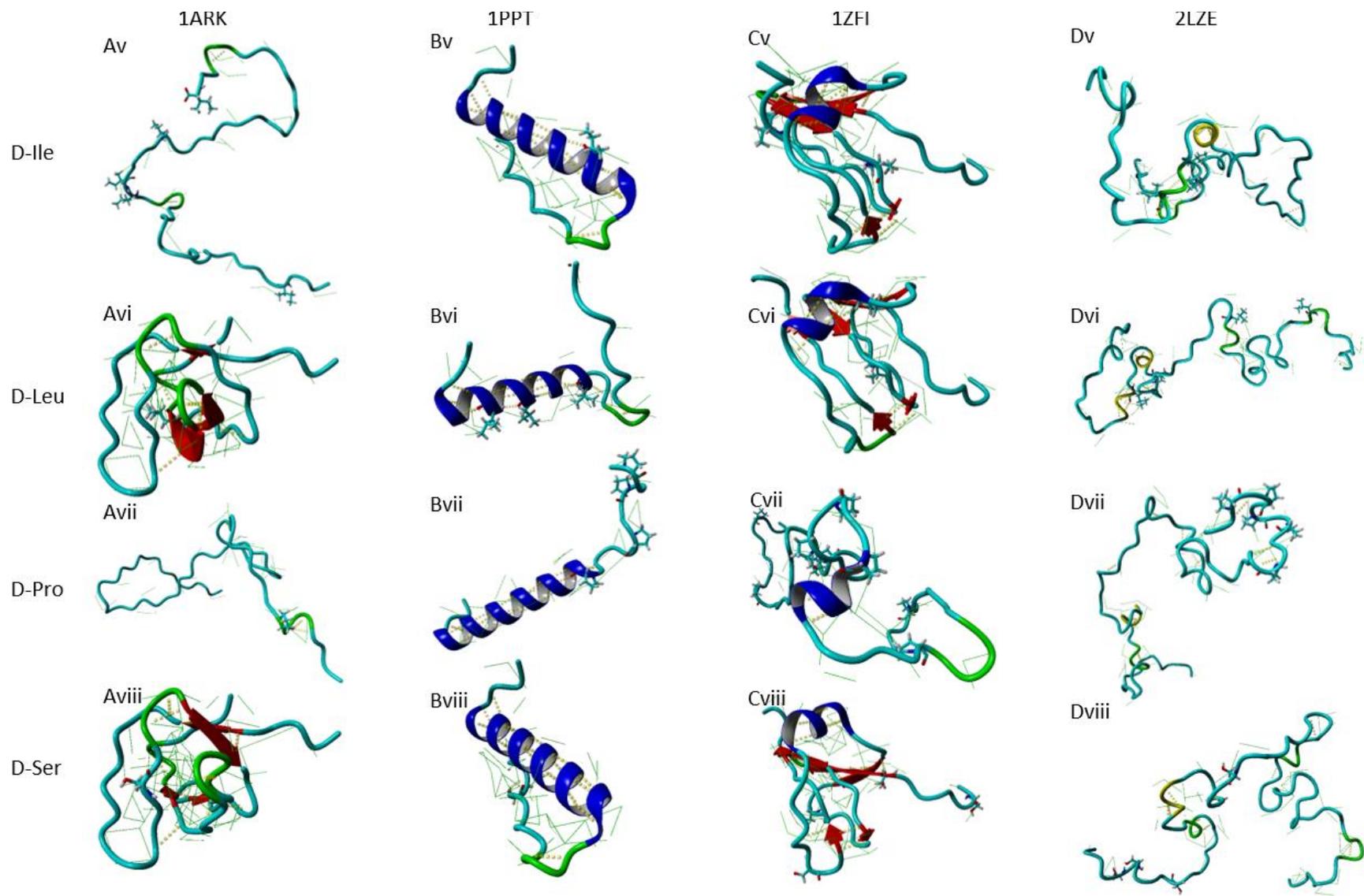


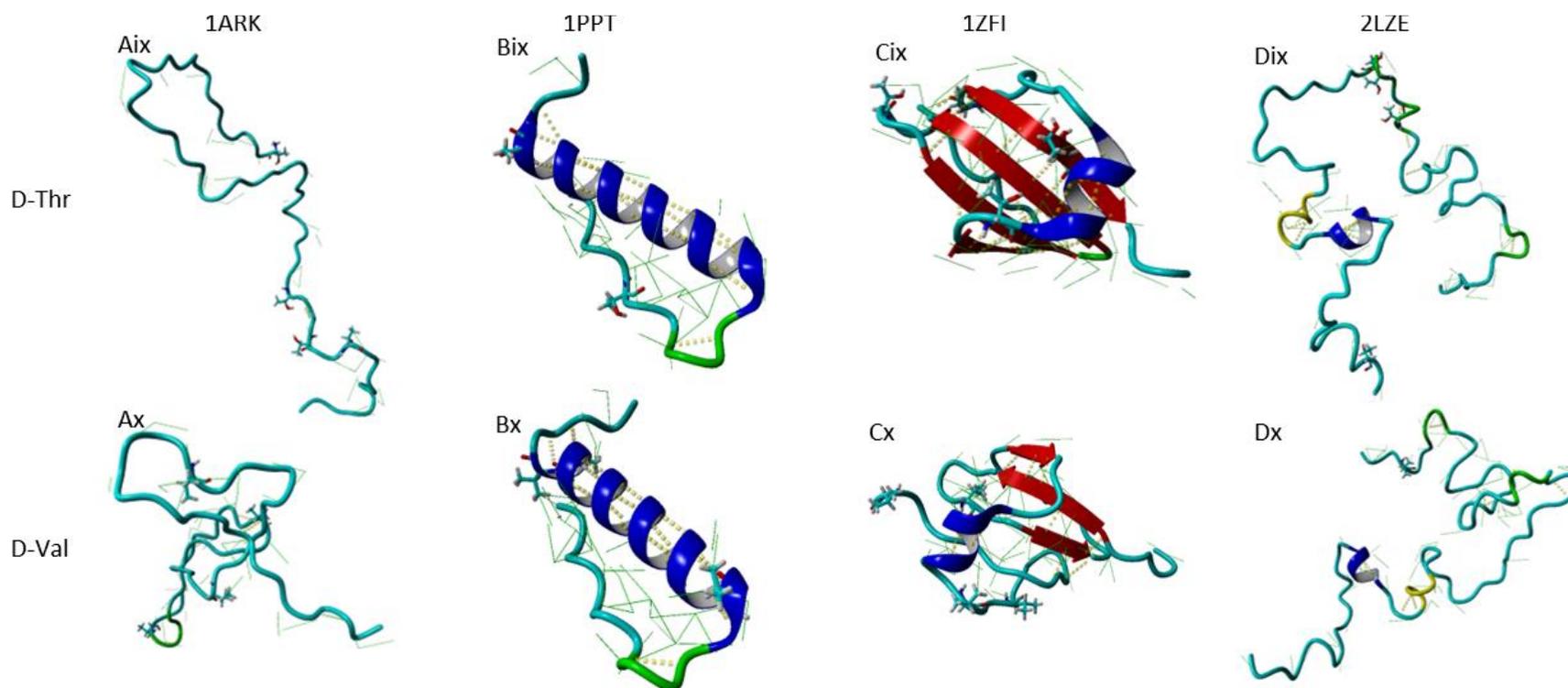




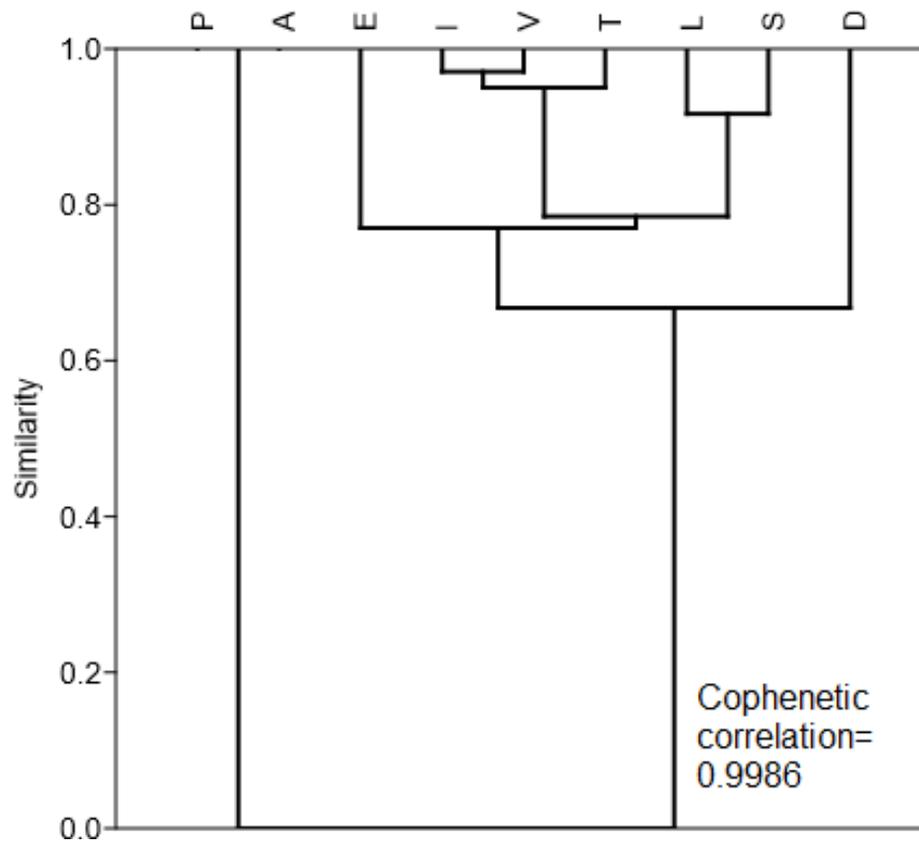
**Figure 1.** 1ARK (Ai-Ax), 1PPT (Bi-Bx), 1ZFI (Ci-Cx), 2LZE (Di-Dx), and their D-amino acid substituted (D-mutated) derivatives after torsion energy minimization (green ribbon, backbone chain with secondary structure depicted; pink dots, atoms in D-mutated residues). Figures A-Di, original polypeptide 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.



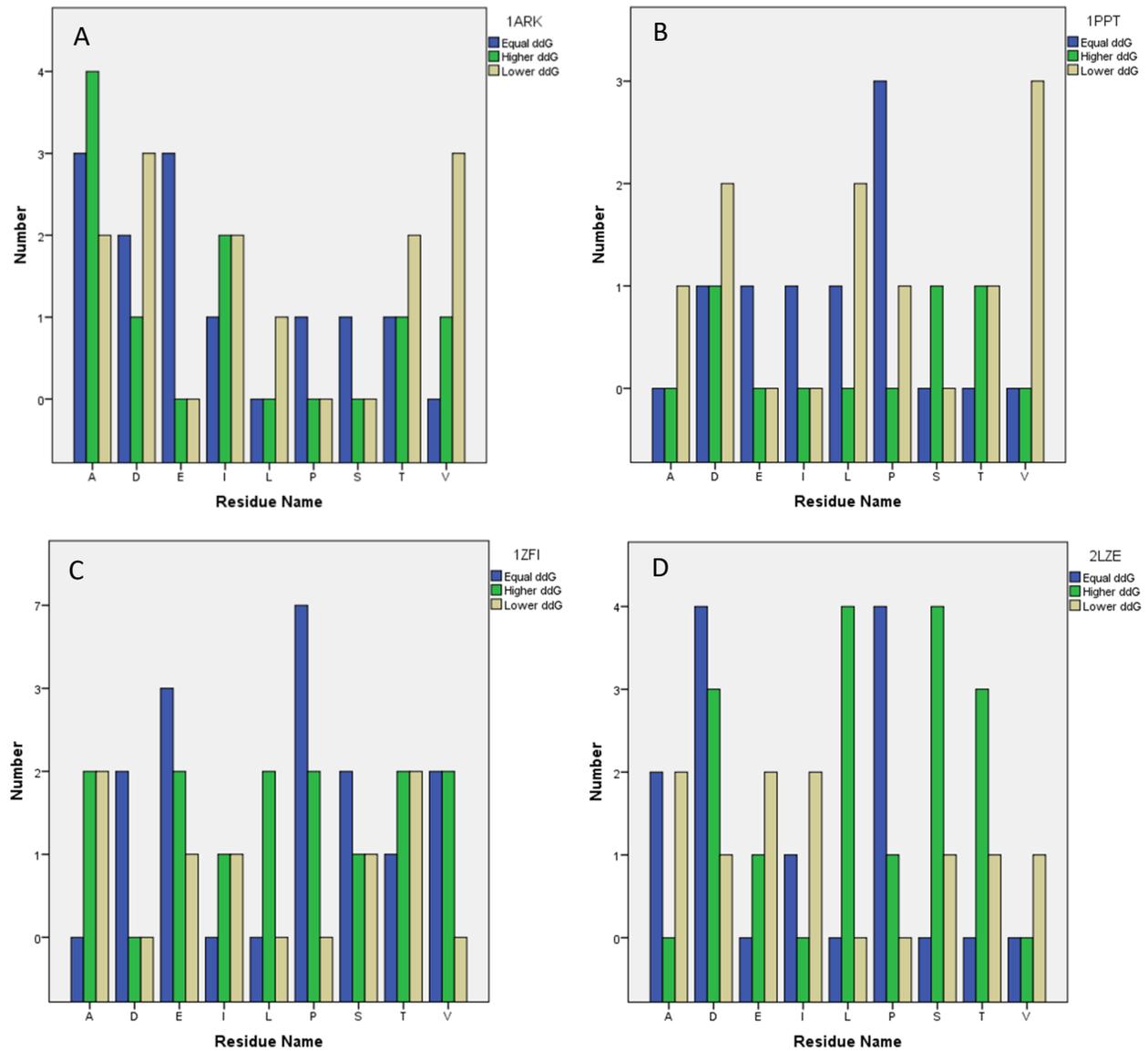




**Figure 2.** Hydrogen bonds and hydrophobic interactions of 1ARK (Ai-Ax), 1PPT (Bi-Bx), 1ZFI (Ci-Cx), 2LZE (Di-Dx), and their partially D-amino acid substituted derivatives after torsion energy minimization (illustrating D-mutated residues in stick style, hydrogen bonds as cylindrical yellow straight 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.



**Figure 3.** Hierarchical clustering (using unweighted pair group method with arithmetic mean algorithm and Bray-Curtis similarity index) dendrogram of 9 chiral prebiotic amino acids within dipeptides.



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

## Tables

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

Polypeptide	Length (residues)	Sequence
<b>1ARK</b>	60	TAGKIFRAMYDYMAADADEVSFKDGDAIINVQAIDEGWMYGTVQRTG RTGMLPANYVEAI
<b>1PPT</b>	36	GPSQPTYPGDDAPVEDLIRFYDNLQQYLNVVTRHRY
<b>1ZFI</b>	67	GSHTPDESFLCYQPDQVCCFCIRGAAPLPSEGECPHPTAPWCREGAV EWVPYSTGQCRTTCIPYVE
<b>2LZE</b>	87	MGAPVPYPDPLEPRGGKHICAICGNNNAEDYKHTDMDLTYTDRDYKNC ESYHKCS DLCQYCRYQKDLAIHHQH HHGGSMGMSGGTGY

**Table 2.** Distribution of prebiotic residues of interest in 1ARK, 1PPT, 1ZFI, and 2LZE peptides.

Class	Residue	1ARK	1PPT	1ZFI	2LZE
<b>Hydrophobic</b>	<b>Ala</b>	9	1	4	4
	<b>Ile</b>	5	1	2	3
	<b>Leu</b>	1	3	2	4
	<b>Pro</b>	1	4	9	5
	<b>Val</b>	4	3	4	1
<b>Acidic</b>	<b>Asp</b>	6	4	2	8
	<b>Glu</b>	3	1	6	3
<b>Hydroxyl</b>	<b>Ser</b>	1	1	4	5
	<b>Thr</b>	4	2	5	4

**Table 3.** Characteristic secondary structure composition in 1ARK, 1PPT, 1ZFI, and 2LZE.

Secondary structure	1ARK	1PPT	1ZFI	2LZE
<b><math>\alpha</math>-Helix (%)</b>	0.0	55.6	9.4	6.9
<b><math>\beta</math>-Sheet (%)</b>	21.3	0.0	44.3	0.0
<b><math>\beta</math>-Turn (%)</b>	15.5	11.1	6.0	4.6
<b>Coil or loop (%)</b>	63.2	33.3	40.3	82.8
<b>3-10 Helix (%)</b>	0.0	0.0	0.0	5.7

**Table 4.** Effects of D-mutations of each species of prebiotic amino acid on normalized stability difference determined using the Cartesian energy minimization approach.

Amino acid D-mutated:		Null	A	D	E	I	L	P	S	T	V
<b>1ARK</b>	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
<b>1PPT</b>	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
<b>1ZFI</b>	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
<b>2LZE</b>	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.

**Table 5.** Secondary structure occupancies, number of mutations, B-factors, and stability of 4 prebiotic polypeptides. Polypeptides substituted with one or more D-amino acids were designated using lowercase letters.

Peptide	Number of mutations	Helix (%)	Sheet (%)	Turn (%)	Coil or loop (%)	3-10 Helix (%)	B-factor	Stability (kcal/mol)
1ARK	0	0.0	21.3	15.5	63.2	0.0	1.0	92.66
1ark_dA	9	0.0	0.0	20.0	80.0	0.0	7.8	347.18
1ark_dD	6	0.0	18.3	18.3	63.3	0.0	7.6	125.22
1ark_dE	3	0.0	26.7	16.7	56.7	0.0	8.4	74.96
1ark_dIdL	6	0.0	0.0	0.0	100.0	0.0	7.7	286.51
1ark_dP	1	0.0	3.3	20.0	76.7	0.0	8.7	109.13
1ark_dS	1	0.0	11.7	18.3	70.0	0.0	8.4	115.70
1ark_dT	4	0.0	0.0	0.0	100.0	0.0	7.4	113.13
1ark_dV	4	0.0	0.0	6.7	93.3	0.0	7.6	236.26
1PPT	0	55.6	0.0	11.1	33.3	0.0	0.0	20.28
1ppt_dA	1	55.6	0.0	11.1	33.3	0.0	7.3	20.19
1ppt_dD	4	55.6	0.0	11.1	33.3	0.0	7.2	34.98
1ppt_dE	1	47.2	0.0	11.1	41.7	0.0	7.3	21.88
1ppt_dIdL	4	50.0	0.0	11.1	38.9	0.0	7.4	46.56
1ppt_dP	4	55.6	0.0	11.1	33.3	0.0	7.6	20.63
1ppt_dS	1	55.6	0.0	11.1	33.3	0.0	7.2	21.13
1ppt_dT	2	58.3	0.0	11.1	30.6	0.0	7.1	24.71
1ppt_dV	3	52.8	0.0	11.1	36.1	0.0	7.0	24.49
1ZFI	0	9.4	44.3	6.0	40.3	0.0	0.0	113.23
1zfi_dA	4	9.0	37.3	6.0	47.8	0.0	8.1	144.46
1zfi_dD	2	0.0	14.9	9.0	70.1	6.0	9.0	122.73
1zfi_dE	6	0.0	29.9	9.0	61.2	0.0	8.2	109.99
1zfi_dIdL	4	0.0	11.9	14.9	73.1	0.0	7.7	128.88
1zfi_dP	9	9.0	23.9	9.0	58.2	0.0	8.2	86.63
1zfi_dS	4	11.9	23.9	3.0	61.2	0.0	8.3	95.54
1zfi_dT	5	11.9	38.8	3.0	46.3	0.0	7.3	133.63
1zfi_dV	4	9.0	16.4	0.0	74.6	0.0	8.3	108.51

**Table 6.** Effects of D-mutations to residue clusters on normalized stability difference and RMSD by using torsion energy minimization approach compared with original all L-amino acid polypeptides (prebiotic polypeptides 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
<b>One cluster</b>												
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#
<b>Two clusters</b>												
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
<b>Three clusters</b>												
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
<b>Four clusters</b>												
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
<b>Five clusters</b>												
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

**Table 7.** 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 signify the number of the following residue that was transformed from L- to D-configuration.

	Number of mutated clusters	$\alpha$ -Helix	$\beta$ -Sheet	$\beta$ -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
All 4-cluster combined 1ppt mutants	4	43.9	0.0	6.7	5.0
All 3-cluster combined 1ppt mutants	3	45.0	0.0	6.7	5.8
All 2-cluster combined 1ppt mutants	2	49.7	0.0	6.7	1.4
All 1-cluster combined 1ppt mutants	1	53.3	0.0	11.1	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