Comparative Simulative Analysis and Design of Single-chain Self-assembled Protein Cages

Fei Xiao,[#a,b] *Longfei Luo,*[#a] *Xin Liu,*[#c] *Ajasja Ljubetič,*[d] *Nengzhi Jin,*[e] *Roman Jerala*[*d] *and Guang Hu*[*a,b]

[a.] MOE Key Laboratory of Geriatric Diseases and Immunology, Suzhou Key Laboratory of Pathogen Bioscience and Anti-infective Medicine, Department of Bioinformatics, Center for Systems Biology, School of Biology and Basic Medical Sciences, Suzhou Medical College of Soochow University, Suzhou, 215213, China

[b.] Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Soochow University, Suzhou 215123, China

[c.] Institute of Blood and Marrow Transplantation, Medical College of Soochow University, Jiangsu Institute of Hematology, The first Affiliated Hospital of Soochow University, Collaborative Innovation Center of Hematology, National Clinical Research Center for Hematologic Diseases, Soochow University, Suzhou 215123, China.

[d.] Department of Synthetic Biology and Immunology, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia.

[e.] Key Laboratory of Advanced Computing of Gansu Province, Gansu Computing Center Lanzhou, 730030，China

* Correspondence: Roman Jerala; roman.jerala@ki.si, Guang Hu; huguang@suda.edu.cn

#These authors contributed equally to this work.

ABSTRACT

Coiled-coil protein origami (CCPO) is a modular strategy for the de novo design of polypeptide nanostructures. It represents a type of modular design based on pairwise-interacting Coiled-coil (CC) units, with a single-chain protein programmed to fold into a polyhedral cage. However, the mechanisms underlying the self-assembly of the protein tetrahedron are still not fully understood. In the present study, eighteen CCPO cages with three different topologies were modeled in silico. Then, molecular dynamics simulations and CC parameters were calculated to characterize the dynamic properties of protein tetrahedral cages at both the local and global levels. Furthermore, a deformed CC unit was redesigned, and the stability of the new cage was significantly improved.

1. INTRODUCTION

Natural biopolymers, including DNA and proteins, are able to self-assemble into well-ordered nanostructures. Based on its characteristic base pairing, DNA has long been used as a programmable building block for the design and self-assembly of complex nanocages.[1] Diverse three-dimensional nanostructures have been made from DNA by using different design strategies,[2] such as a one-pot approach,[3] hierarchical self-assembly,[4] and the DNA origami technique.[5] Owing

to numerous weak cooperative interactions, the rational design of new protein folds has been very challenging[6]. The symmetry-based method was proposed for self-assembling protein cages by fusing natural protein oligomerization domains and designing different protein-protein interfaces.[7] To date, several protein cages have been experimentally characterized, including tetrahedra,[8-9] cubes,[10] octahedra,[11] and icosahedra.[12] However, this strategy has a limitation in that only symmetric structures can be synthesized. Recent advances in application of generative machine learning enabled the de novo design of new compact protein folds and diverse assemblies.[13-16]. Nevertheless, due to the encapsulation property of the inner large cavity, DNA and protein cages with shapes not found in natural proteins and therefore difficult to predict by ML-based methods might have potential applications, from drug delivery to nanomaterials and as a reaction chamber for enzymatic reactions.[17-19]

Inspired by DNA nanotechnology,[20] a modular approach was developed to assemble protein polyhedral cages from a single-chain polypeptide.[21] In this method, orthogonal dimerizing coiled-coil segments were used as interacting modules to construct the uniquely defined topological protein folds.[22] Coiled–coil dimers are ubiquitous protein segments whose properties as structural elements in synthetic biology have been investigated.[23] Unlike DNA antiparallel duplexes, coiled-coil dimers may form in either a parallel or an antiparallel orientation. The major advantage of the modular approach is that protein cages can fold spontaneously without the need for mixing, denaturing, and slowly annealing multiple polypeptide chains.[24] Resembling in many aspects of DNA origami, a modular approach based on a coiled-coil complementarity polypeptide can be considered protein origami (coiled-coil protein origami, CCPO),[25] which represents the first steps of this technological platform. We have so far explored only a minuscule fraction of the features and potentials of this new domain.

In this respect, theoretical approaches hold much potential for providing valuable insight into their structures and self-assembly. Two mathematical models, namely, polyhedral links[26-28] and a strong trace model,[29] were introduced for the prediction of complex topologies of DNA and protein nanocages. On the other hand, molecular dynamics (MD) simulations coupled with principal component analysis (PCA) analysis have been successfully used to characterize the structure and dynamics of DNA nanocages,[30-31] compare the stabilities of octahedral DNA nanocages with different structural properties,[32-36] and characterize the opening/closing molecular mechanism of a DNA truncated octahedron[37-38] to investigate the self-assembly of DNA origami-based assemblies.[39] Using oxDNA,[40] a coarse-grained model was proposed, which has been applied to characterize the structures of DNA star-tile-based nanocages[41] and the self-assembly of DNA origami[42] and single chain knotted DNA nanostructures.[43] Recent high-resolution determination of a CCPO triangle confirmed the structure according to the design.[44] While the difficulties in determination of high-resolution structure of CCPOs suggested that the flexibility of linkers may introduce high flexibility, stronger evidence would could be provided by a simulation.

In this work, we investigated the stability and dynamic properties of CCPO cages based on MD simulations. First, eighteen protein cages in the shape of tetrahedra, square pyramids and triangular prisms were modeled in silico, and their stability and dynamical properties were investigated using the MD approach. Then, the stability and dynamical properties of two CCPO cages with the same topology and different initial conformation were compared. The parameter analysis and structural analysis of all the constituent CC blocks were performed, and a pair of CC blocks were found with the weakest stability. We further mutated and redesigned this CC dimer to improve its stability of the redesigned cage. Our computational work provides detailed structural insight as well as

dynamical properties of single-chain protein tetrahedral cages and hence aids rational design in new protein cages based on the modular approach.

2. MATERIALS AND METHODS

2.1. Building protocol

We have followed the experimental framework of Ljubetič et al. for the design and sequences of amino acids comprising a series of CCPO structures[21, 45] in the shape of tetrahedra, square pyramids and triangular prisms. The atomistic model of the protein cages is constructed through the following three steps using the CoCoPOD platform in a semiautomated manner[45]. The design platform CoCoPOD provided a suitable environment for the design of different polyhedra, showcasing the utility of the developed software.

Three representative topologies, tetrahedron, square pyramid and triangular prism, formed by 12, 16 and 18 CC segments, respectively, were confirmed by both small angle X-ray scattering (SAXS) and single particle TEM reconstruction. The density maps have been deposited in the EMDB database under the following IDs: EMD-3781 (TET12SN), EMD-3788 (PYR16SN), and EMD-3789 (TRIP18SN). We investigated eighteen structures of the abovementioned shapes, including fifteen cages that are best-fit to the SAXS data and three cages with the largest volume (Table S1). To better understand the characteristics of the systems, we provide $TET12_{1.10}SN\text{-}f_5$ as an example for illustrating the naming conventions, which contains all information for polyhedral structures. The first three letters are the polyhedron type (TET = tetrahedron; PYR = pyramid; TRIP = trigonal prism), followed by the number of CC segments. The subscript denotes the topology and circular permutation of each polyhedron. The next labels denote the type of CC modules used (S = soluble; SN = soluble, negatively charged), linker type (f = flexible; c =

charged), and, in subscript form, the length of the linker. In cases where two variants have the same name (e.g., different ordering of CC modules), the letters b, c, d, and so forth are appended.[45]

2.2. **Simulation methodology**

All-atom molecular dynamics simulations were performed using GROMACS 5.1.4.[46] The bonded and nonbonded description of the interactions between the various atoms has been described using the Amber99SB-ILDN force field.[47] The initial structures of CCPOs were constructed using the CoCoPOD platform. Each structure was immersed in a triclinic box filled with TIP3P water molecules, imposing a minimum distance of 12 Å between the solute and the box. Initially, we performed a series of energy minimization steps to eliminate any bad contacts in the initially built structures. Each minimization step involves 5000 steps of steepest descent followed by 5000 steps of the l-bfgs method. After the energy minimization, the equilibration was conducted in three steps of 100 ps each, and the system was heated from 10 K (step 1) to 150 K (step 2) and finally to 300 K (step 3). Finally, 200 ns NPT production simulations are performed at 300 K and 1 atm pressure with a 2 fs integration time step. We have implemented periodic boundary conditions across the system using a truncated octahedral cell. We use particle mesh Ewald (PME) techniques integrated with the Gromacs package to account for the long-range part of the electrostatic interactions.[48-49] During the dynamics, all the bonds involving hydrogen are restrained using the SHAKE algorithm.[50-51] The Langevin thermostat with a collision frequency of 1 ps-1 is used to maintain the constant temperature and pressure. The accelerated GPU version of PMEMD[52-53] was performed on Nvidia GTX1060 series cards. The images and graphics of the structures shown here were generated using the software packages VMD[54] and PyMOL.[55]

2.3. **Structural deviation and atomic fluctuation**

Root mean square deviations (RMSDs) and root mean square fluctuation (RMSFs) of the trajectories were calculated by using the GROMACS 5.1.4 analysis tools. The first 50 ns were removed, and the remaining trajectories were used for the analyses. Two types of RMSFs have been calculated. The local RMSF was calculated for each of the twelve strands extracted from the total cage and singly fitting each of them without considering the total cage motion. The global RMSF was calculated by fitting the total cage conformation. Both RMSFs are calculated over the Cα atoms. The values of global RMSF are influenced by the total cage motion.

2.4. **Surface and Volume analysis**

Manipulation of self-assembling protein modules is the key to achieving highly controllable protein cages for encapsulation and drug delivery. Thus, its molecular surface area and volume are very important to estimate its true loading capacity. To account for the variation in the overall size and volume of the CCPOs during the simulation, we calculated its radius of gyration (Rg) with the GROMACS 5.1.4 analysis tools using only Cα atoms. The computed volume of our model using Voronoi diagram-based geometrical methods to compute the void volume.[57]

2.5. Principal component analysis

The collective motions of the cages were investigated by principal component analysis (PCA). For the cage systems, PCA was performed using the Cα atoms of the MD simulation trajectories through gcovar integrated with Gromacs. The trajectory matrix contains in each column Cartesian coordinates for a given atom at each time step (x(t)). Fitting the coordinate data to a reference structure result in the proper trajectory matrix (X). The trajectory data are then used to generate a covariance matrix (C), the elements of which are defined as:

$$C_{ij}=\langle(x_i-\langle x_i\rangle)(x_j-\langle x_j\rangle)\rangle \quad (1)$$

where $\langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$ denotes an average performed over all the time steps of the trajectory. The principal components (PCs) are obtained by diagonalization of the covariance matrix C.

$$C = V \Lambda V^T \quad (2)$$

This results in a diagonal matrix L containing the eigenvalues as diagonal entries and a matrix V containing the corresponding eigenvectors. If the eigenvectors are sorted such that their eigenvalues are in decreasing order, the eigenvector with the largest eigenvalues (i.e., the first principal component, PC1) accounts for the highest proportion of variance within the data. The second principal component (PC2) is orthogonal to the first one and accounts for the second-highest proportion of variance. The first two components were obtained using gmx anaeig, and the score plot of PC1 vs PC2 was used to evaluate differences in the oscillatory motion of the structures throughout the MD simulations. In addition, the dynamics cross-correlation maps are calculated from the PCA by using the g_covar utility of GROMACS. In this map, each value indicates whether Ca atoms move in the same or opposite direction.

2.6. **SAXS data analysis**

To determine whether the CCPOs maintained their conformations, we compared the computed SAXS curves of the average structures of CCPOs from equilibrated MD simulations to the experimental SAXS data. Their theoretical SAXS profiles were computed with CRYSOL.[58]

2.7. **Parameterization of coiled coils.**

The original mathematical parameterization of coiled coils here is from Crick[59] and has been developed since,[60] including in CCCP,[61] CCBuilder,[62] which are web-based applications for parametric modeling of coiled coils, and ISAMBARD.[63]

3. RESULTS AND DISCUSSION

3.1. Most CCPO cages are stable during simulations

To investigate the structural stability and dynamics of the systems during the whole simulation time, the root-mean-square deviation (RMSDs), gyration radius (Rg), and volume (V) of all CCPO cage simulations with respect to the starting structures were calculated and plotted, as shown in Fig. S1-S4 and Table S1. The initial conformations were taken as the best-fit model to the experimental SAXS scattering curves.[45] These results show that RMSD and Rg values converge, confirming that CCPO cages reached the equilibrium state after 200 ns of simulation, except for $TET12_{1.10}SN\text{-}f_5$, which reached equilibrium after 500 ns of simulation. The RMSDs with respect to their initial minimized structure saturates to values between 0.58 and 1.29 nm for TET12 cages, while $TET12_{1.10}S\text{-}f_5$ with the best conformation has the largest RMSD. For the PYR16 and TRIP18 cages, the RMSDs range from 0.80 to 0.88 nm and from 1.01 to 1.33 nm, respectively. For illustration, we selected three representative systems for comparison and elucidation, as depicted in Fig. 1 and Fig. S1. By superimposing the mean and final structures onto the initial structures (Fig. 1a), as well as the mean structures extracted every 50 ns from MD trajectories were superposed onto the initial structure (Fig. S1a), we observed that the overall shapes of the three systems did not undergone significant changes. In addition, PCA was performed to reveal changes in the slow oscillatory motion of the structures throughout the MD simulations (Fig. 1b). It was observed that within the initial 100 ns of MD, some conformational changes occur in each system, but between 100 ns and 200 ns, the conformations become stable. RMSDs for $TET12_{1.10}S\text{-}c_6$ (blue lines), $PYR16_{2.15R}S\text{-}f_5$ (red lines) and $TRIP18_{7.5R}SN\text{-}f_5$-oblique (green lines) are shown in Fig. S1b. The larger RMSD values for the PYR16 and TRIP18 cages are mainly due to their larger size. The GR, a measure of the structure compactness, shows a similar trend with RMSD, confirming the

stability of CCPO cages. For example, as shown in Fig. S1c, the average GR values settle to ~3.1, ~3.4 and ~3.6 nm for TET12$_{1.10}$S-c$_6$, PYR16$_{2.15R}$S-f$_5$ and TRIP18$_{7.5R}$SN-f$_5$-obl, respectively. It can be seen that the GR values keep decreasing, which indicates that CCPO cages experience slight collapsing during the MD simulations.

To estimate the maximal cargo carrying capacity, the shape and volume of CCPO cages are other important indicators. We calculated the cavity volumes during the simulation to account for the variation in the overall size of CCPO cages. The volumes of most of the CCPO cages attained a dominant decrease during the initial 10 ns of run time and converged in the last 80 ns of simulation. The variations in volumes (ΔV) of all cages are also listed in Table S1. As illustrated in Fig. S1d, the volumes of TET12$_{1.10}$S-c$_6$, PYR16$_{2.15R}$S-f$_5$ and TRIP18$_{7.5R}$SN-f$_5$-obl changed from 54.6 nm^3 to 50.1 nm^3, from 77.3 nm^3 to 74.1 nm^3, and from 125.1 nm^3 to 83.2 nm^3, respectively. Accordingly, the values of ΔV are also reflected in the shape and size of CCPO cages, which means that the shapes of most CCPO cages do not undergo a large deviation or even a collapse and provide the ability to host cargo molecules at the central cavity.

To further confirm the stability of CCPO cages, we also compared our MD simulations with the SAXS data (Fig. S5). For each cage, 10 configurations of the last 100 ns in simulation were extracted from the trajectory with a time step of 0.1 ns, and the Guinier plot was used to compare the simulation and experimental results. As listed in Table S1 and Fig. 1c, the R^2 for the most fitting of Guinier plots approaches 1, which shows that the structures during the simulation show good agreement with the SAXS experimental data. In line, the fitting R^2 of Guinier plots in TET12$_{1.10}$S-c$_6$, PYR16$_{2.15R}$S-f$_5$-obl and TRIP18$_{7.5R}$SN-f$_5$ are 0.97, 0.98, and 0.93, respectively. The good agreement between the SAXS experiment and MD simulation results indicates that most of the CCPO cages maintain their shapes during the simulation.

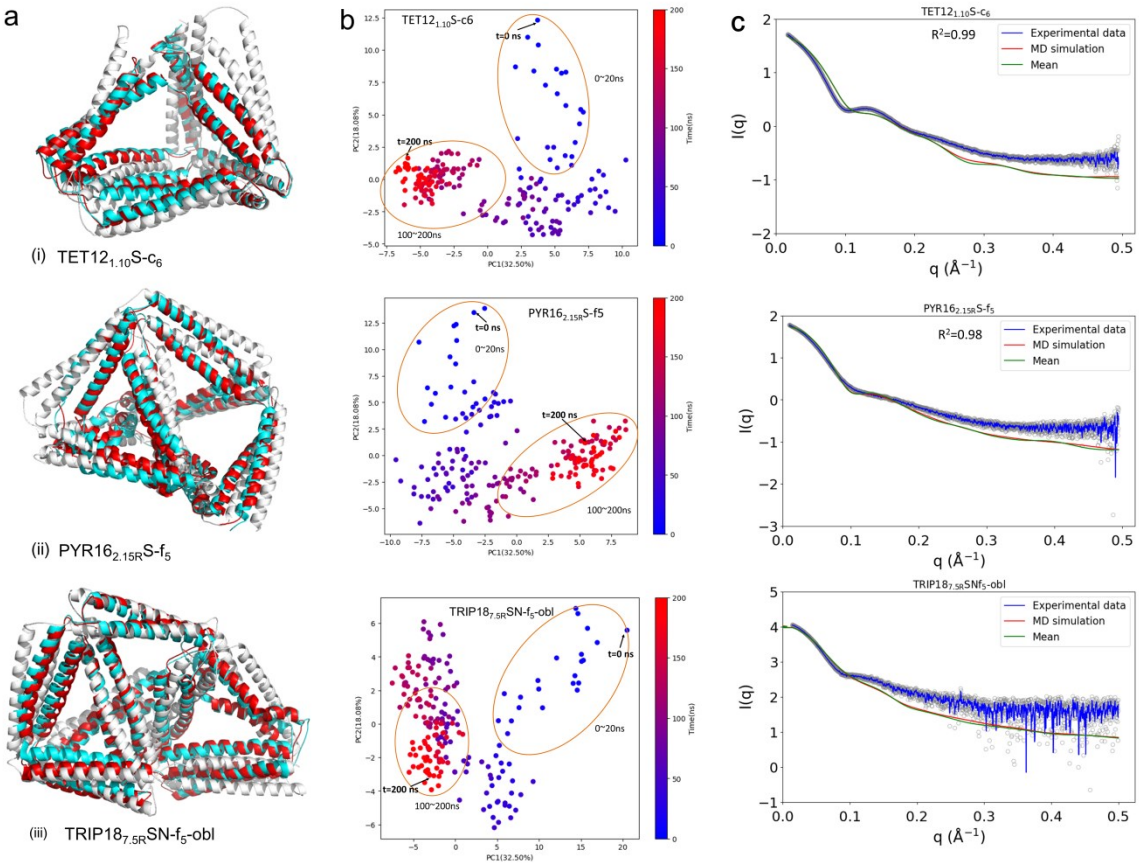


Figure 1. Molecular dynamics results for TET12$_{1.10}$S-c$_6$, PYR16$_{2.15R}$S-f$_5$ and TRIP18$_{7.5R}$SN-f$_5$. (a) The final (red) and mean (cyan) structures of computational models of three CCPO cages were superposed on their corresponding initial conformations (gray). (b) PCA score plot on conformations extracted from MD simulations of TET12$_{1.10}$S-c$_6$, PYR16$_{2.15R}$S-f$_5$ and TRIP18$_{7.5R}$SN-f$_5$-obl (up to down). The points representing the initial and final conformations are indicated in the figure with arrows. The color of the points corresponds to the transition from blue to red, representing the conformations at times ranging from t=0ns to t=200ns, respectively. (c) Comparison of the MD simulation results and the experimental SAXS data. Blue lines with gray standard derivation are calculated from experimental SAXS data, and red and green lines are scattering intensity calculated for the MD simulations and the mean structure, representatively.

3.2. Dynamic comparison between $TET12_{1.10}SN\text{-}f_5^{best}$ and $TET12_{1.10}SN\text{-}f_5^{largest}$

In the previous section, we showed that different variants of CCPO cages share similar dynamic behavior. Among our studied systems, the two $TET12_{1.10}SN\text{-}f_5$ cages shared the same topology and composition but with different initial conformations and were named "$TET12_{1.10}SN\text{-}f_5^{best}$", whose conformation was obtained by fitting with the best-fit-to-SAXS data, while "$TET12_{1.10}SN\text{-}f_5^{largest}$" had the largest volume.[45] It would be interesting and useful to study if the cages with different initial conformations have effects on their dynamics and stabilities.

Overall, the detailed dynamic analysis of CCPO cages found that $TET12_{1.10}SN\text{-}f_5^{best}$ showed the largest RMSD, as well as relatively large Rg and ΔV. The instability of $TET12_{1.10}SN\text{-}f_5^{best}$ is likely due to the most deformed initial structures. As shown in Fig. 2 (a), the RMSD with respect to their initial minimized structure saturates to values of 0.91 nm and 0.80 nm for $TET12_{1.10}SN\text{-}f_5^{best}$ and $TET12_{1.10}SN\text{-}f_5^{largest}$, respectively. The Rg values with respect to their initial minimized structure settle from 3.29 to 2.97 nm and 3.30 to 3.13 for $TET12_{1.10}SN\text{-}f_5^{best}$ and $TET12_{1.10}SN\text{-}f_5^{largest}$, respectively. The ΔV along the MD simulation for $TET12_{1.10}SN\text{-}f_5^{best}$ is 15.21 nm, which has been shifted to 3.50 for $TET12_{1.10}SN\text{-}f_5^{largest}$. The much smaller RMSD and ΔV values of $TET12_{1.10}SN\text{-}f_5$ are the largest and are more stable than those of $TET12_{1.10}SN\text{-}f_5^{best}$. These general results indicate that the initial conformations of CCPO cages have a large effect on their global conformational dynamics, which has also been confirmed by the low overlap values between PCA modes (Fig. S6).

To investigate the effect of the initial conformation on the dynamics of CCPO cages, we compared the conformational dynamics between $TET12_{1.10}SN\text{-}f_5^{best}$ (Movie S1) and $TET12_{1.10}SN\text{-}f_5^{largest}$ (Movie S2). It is easy to see that the $TET12_{1.10}SN\text{-}f_5^{largest}$ cage remains stable during the MD simulation, while the $TET12_{1.10}SN\text{-}f_5^{best}$ cage demonstrates remarkable deformation. After

approximately 80 ns, the initial structure of TET12$_{1.10}$SN-f$_5^{best}$ started a transition to a distortion conformation, which is characterized by a further movement of the CC segment of APHshSN. Until 200 ns, the end of APHshSN interacts with another CC segment of P5SN-P6SN through two hydrogen bond networks between GLY196 or LEU194 in APHshSN and GLN312 in P5SN (see Fig. 2b and Movie S1). Then, we used two residue pairs, GLY196-GLN312 (green sticks) and LEU194-GLN312 (yellow sticks), as references to quantify the movement of APHshSN. The distance between the first pair changed from 47.1 Å (green arrows in B) to 3.6 Å (red arrows in B), and the second pair changed from 52.5 Å to 3.4 Å. The final distance between two residue pairs represents the distances required to form strong hydrogen bonds, which was maintained after 200 ns. Some deviation occurs in the position of the oscillations after 0.40 nm and is likely due to the distorted structure of the TET12$_{1.10}$SN-f$_5^{best}$ cage after 80 ns of MD simulation.

Fig. 2c shows the cross-correlation maps for TET12$_{1.10}$SN-f$_5$ with two different conformations. The lower triangular part of the map shows cross-correlations for the TET12$_{1.10}$SN-f$_5$ ^{best} cage, while the upper triangular part shows cross-correlations for the TET12$_{1.10}$SN-f$_5$ ^{largest} cage. In the map, values range from -1 to 1, in which the red regions with positive values represent residues moving in the same direction, and the blue regions with negative values denote residues moving in the opposite direction. The correlation map of both cages shows six positively correlated elements (red blocks) corresponding to strong coupling within each of the six coiled-coil segments. However, the TET12$_{1.10}$SN-f$_5^{best}$ cage (lower triangular part) shows larger negative correlations than the TET12$_{1.10}$SN-f$_5^{largest}$ cage (upper triangular part), which means that the structure of the TET12$_{1.10}$SN-f$_5^{best}$ cage is contracting more than the TET12$_{1.10}$SN-f$_5^{largest}$ cage during simulation. In particular, four square boxes in the lower triangular part of the map highlight the negative correlations between the APHshSN dimer and the P5SN-P6SN dimer.

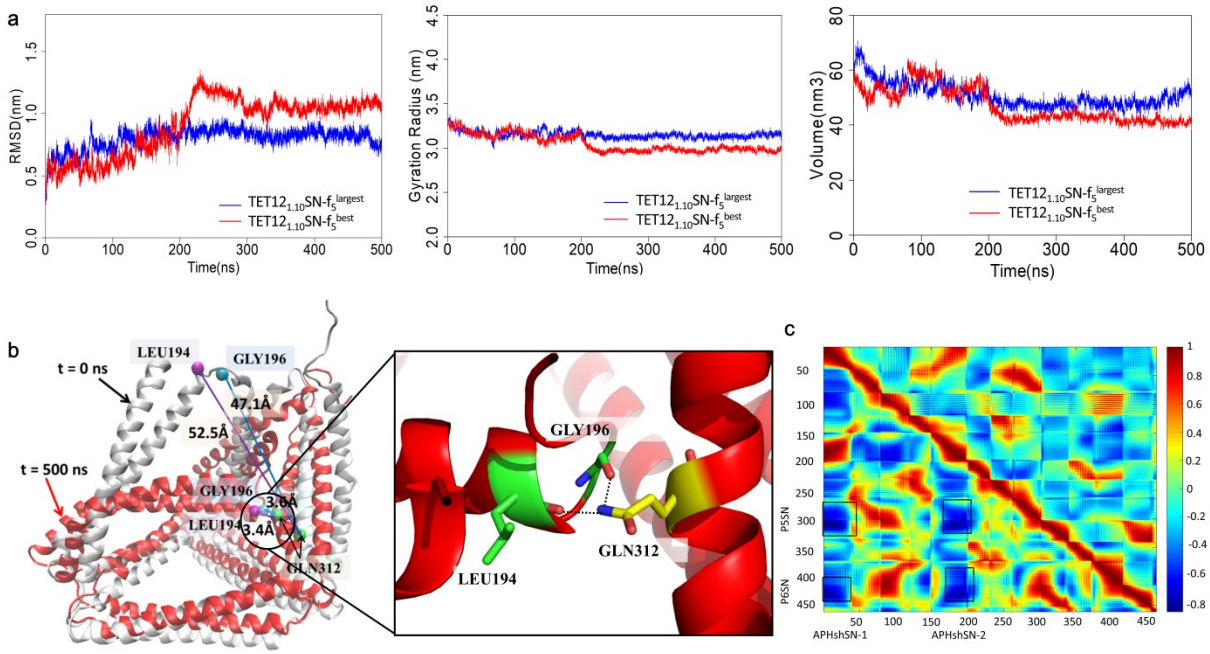


Figure 2. Dynamic results for the comparison between TET12$_{1.10}$SN-f$_5^{best}$ and TET12$_{1.10}$SN-f$_5^{largest}$. (a) Time evolution of the RMSD, the radius of gyration (Rg), and the volume over the entire trajectory for TET12$_{1.10}$SN-f$_5^{best}$ (red line) and TET12$_{1.10}$SN-f$_5^{largest}$ (blue line). (b) The structural deformation (red color) of the TET12$_{1.10}$SN-f$_5^{best}$ cage after superposition on its corresponding initial conformation (green color). (c) The dynamical cross-correlation map for TET12$_{1.10}$SN-f$_5^{best}$ (lower triangular part) and TET12$_{1.10}$SN-f$_5^{largest}$ (upper triangular part). Four square boxes highlight the negative correlations between the APHshSN dimer and the P5SN-P6SN dimer.

3.3. APH4SN in the TET12$_{2.3}$SN-f$_5$b cage shares a large deformation

In this section, root-mean-square fluctuations (RMSFs) were calculated to describe the conformation mobility and dynamic properties of each coiled-coil segment of CCPO cages, as shown in Fig. S7. To compare the stability of the building blocks, only RMSFs of coiled-coil

segments were calculated, while the RMSFs values of linkers are not shown. Although we found that the overall structure of TET12$_{2.3}$SN-f$_5$b is stable, prominent variations were observed for the APH4SN segment. In Fig. 4a, the RMSFs for APH4SN show much larger values (0.22±0.6 nm and 0.22±0.7 nm for two APH4SN strands, respectively), whereas for other segments, they share similar smaller average values that vary from 0.07 nm to 0.1 nm.

The large deformation of APH4sn in the TET12$_{2.3}$SN-f$_5$b cage can be further explained by 3D conformation fitting (Fig. 3b) and PCA motions (Fig. 3c, d). Superimposition of the initial (yellow) and final (magenta) structures of TET12$_{2.3}$SN-f$_5$b are shown with cartoon representations. The coiled-coil dimer APH4SN, showing the largest conformational change, is outlined by the black frame. Structural alignment of the start and end TET12$_{2.3}$SN-f$_5$b structure displays an average RMSD of 0.88 nm, mainly caused by the large displacement of the middle region of the APH4SN segment. As shown in Movie S3, during the MD simulation, amino acids around the middle of the APH4SN segment displayed obvious variations, in agreement with the 3D conformation fitting.

In addition, the largest fluctuation of APH4SN in the TET12$_{2.3}$SN-f$_5$b cage was further confirmed by the PCA motion of the MD trajectory. The PCA method is designed to detect the dominant internal motions among the various components of the systems by a small set of PCA modes (PCs). For the TET12$_{2.3}$SN-f$_5$b cage, the first and second PCA modes account for 34% and 20% of the total variance, respectively, while the motions of these two modes are shown in Fig. 3c and d. Both PC1 and PC2 show the breathing motion of the cage, which corresponds to the expansion/contraction of the interior cavity within the polyhedral cage. In particular, the first mode shows the large vibration of three linkers, while the second mode shows the large deformation of the APH4SN segment. Accordingly, PC2 describes the determinant motions of the large deformation observed in the TET12$_{2.3}$SN-f$_5$b cage.

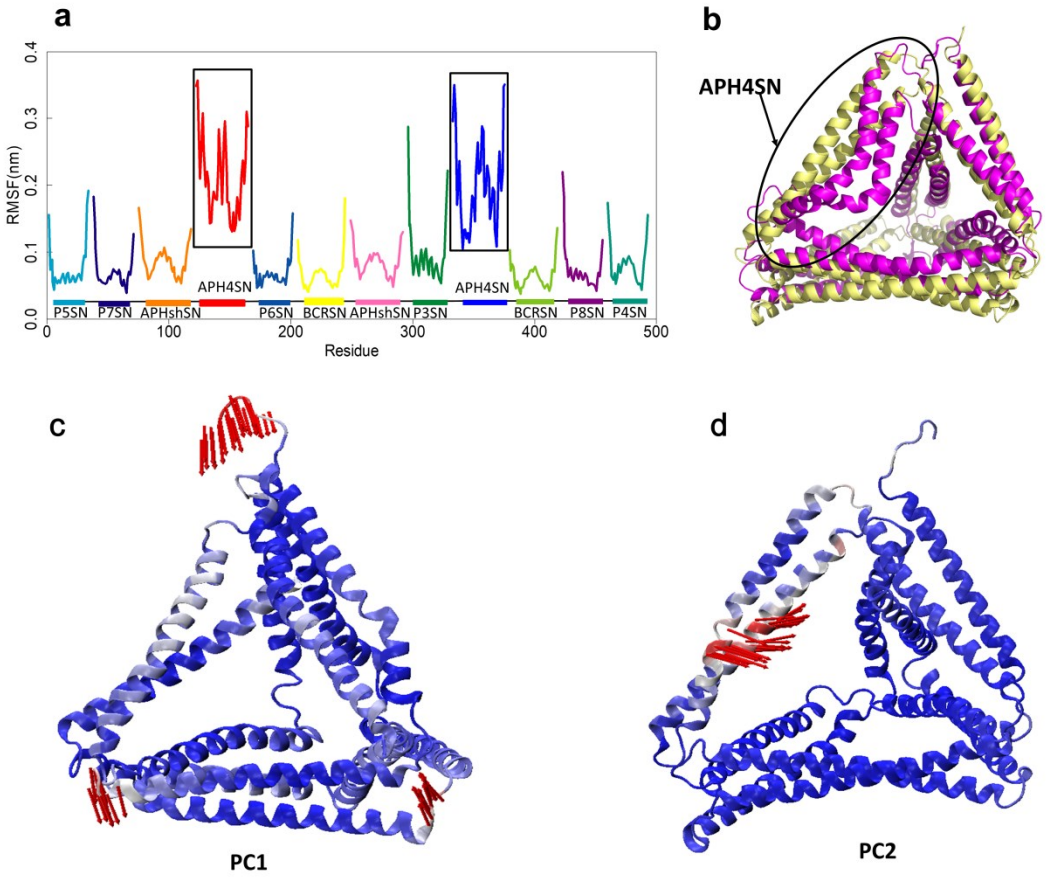


Figure 3. (a) Per Cα residue RMSF for each strand of TET12$_{2.3}$SN-f$_5$b, calculated over the simulation time. Each strand is represented with a different color. (b) Superimposition of the initial (gray) and final (colored) structures of TET12$_{2.3}$SN-f$_5$b with cartoon representations. APH4SN shows a large conformational change, which is highlighted by a black circle. (c) PC1 describes the large vibration of three linkers, while (d) PC2 describes the largest motion for the particular regain in the APH4SN, shown by the red arrows.

3.4. Structural parameters and helix structures of CC dimers

In CCPO cages, CC dimers are regular and relatively rigid building blocks, which are stabilized by a characteristic 'knobs-into-holes' packing.[64] The geometry and packing mode of CCs can be

described by Crick equations that involve a set of structural parameters. We calculated eight structural parameters for six CC dimers based on the 200 ns TET12$_{2.3}$SN-f$_5$b MD trajectory (Fig. S8), including the pitch and radius of the superhelix, residues per turn, phi-, psi-, and alpha-angles, crick angle and its deviation.[63] The mean values along the MD simulation, with their standard deviations of six CC dimers, are reported in Table S2. These geometrical parameters of most CC dimers show similar values and few deviations along the MD trajectory, indicating good stability and a regular geometry of those CC dimers that is maintained over the entire simulation time. Among these structural parameters, the alpha-angles (α) show large variation for APH4sn in comparison with other CC dimers, suggesting that α provides a useful matrix to analyze the possible geometrical deformation of APH4sn. Fig. 4 (a) shows the alpha angles calculated as a function of time for each CC dimer in TET12$_{2.3}$SN-f$_5$b. The mean α value of APH4sn is approximately 16.76, while the values for other CC dimers range between 9.45 and 12.5. The alpha-angle (α) is the angle that the peptide plane makes with the helix axis, with a value of approximately 13 degrees. The larger alpha-angle of APH4sn describes a different hydrophobic core packing geometry, which will affect the global helix axial rotation in structures.

Moreover, secondary structure analysis of the helical structure of coiled-coil segments in TET12$_{2.3}$SN-f$_5$b was performed over simulation time to quantify and compare the deviation from their canonical conformation. The detailed values of the share of the secondary structure content of each segment in TET12$_{2.3}$SN-f$_5$b are listed in Table S3. We found that, except for APH4SN with large deviations in helical parameters, the overall secondary structure was maintained during the simulation. The secondary structure analysis of APH4SN is shown in Fig. 4b, which demonstrates the formation of defects in the APH4SN structure. The APH4SN segment is

composed of a helical structure at the beginning of the MD simulation and exhibits a transition between the β-turn and α-helix of the secondary structure.

By investigating the CC parameters and secondary structure component, the structural determinant and deformation mechanism of APH4sn in TET12$_{2.3}$SN-f$_5$b are proposed. The analysis of the simulation of the APH4sn dimeric coiled coil reveals progressive helix unfolding, which is determined by the packing model with a larger alpha-angle. With the goal of using CCs as nanomechanical building blocks, the helix propensity and hydrophobic core packing provide two useful matrixes for the parametric molecular design.

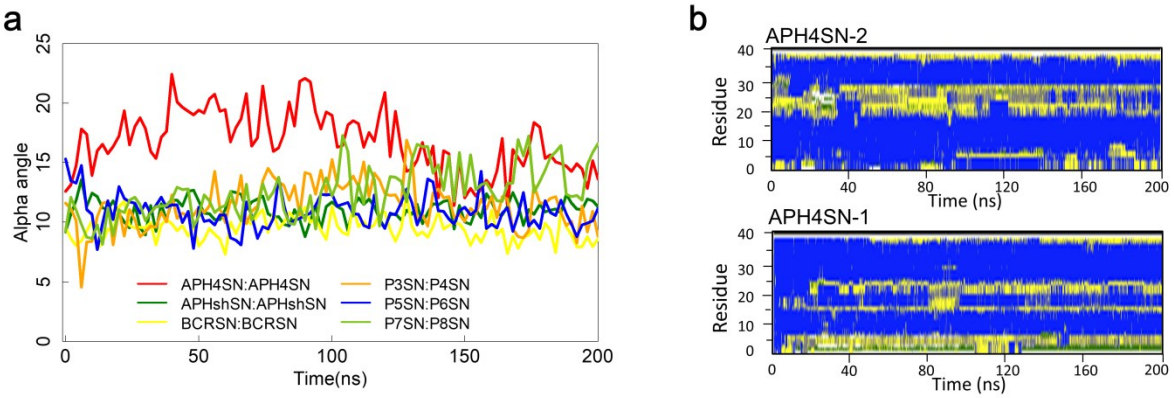


Figure 4. (a) Alpha angle parameter for six coiled-coil segments of TET12$_{2.3}$SN-f$_5$b, calculated over the simulation time. (b) Secondary structure assessment over time. The blue and yellow colors represent the helix structure and turn for each amino acid during the MD simulation.

3.5. Implications for the Molecular Design

The primary sequence of coiled coils is characterized by a repetitive pattern of seven amino acid residues, called heptad repeats (abcdefg)$_n$.[65] Residues at the *a* and *d* positions are predominantly hydrophobic, whereas residues at the *e* and *g* positions are frequently charged and can participate in interhelical Coulombic interactions. The binding stability and specificity are mainly determined

by the burial of hydrophobic side chains at positions *a* and *d* and complementary electrostatic interactions between residues at positions e and g. Utilizing this simple design, CCs serve as model systems for studying protein folding and stability and design of orthogonal building modules.[66-68] Even minor differences in the binding interface (e.g., β-branching side chains or Asn at position a or d) can therefore dramatically affect the CC structure. This suggests that a molecular design strategy relies on the mutation of a single residue.

The previous results demonstrate that APH4SN in TET12$_{2.3}$SN-f$_5$b is not stable, suggesting that APH4SN is the key design parameter for tuning a more stable TET12$_{2.3}$SN-f$_5$b cage. When mapping the structures of the APH4SN observed in the MD simulation into such helical-wheel diagrams, the deformation region was assigned to the heptad including two lysine residues at the *a* and *d* (*a'* and *d'*) positions (Fig. 5a). By analyzing the structural deformation region, we suggest that the residues Lys (K) at positions a and d' are likely to contribute to the deformation of APH4SN. This sequence was based on the designed set of orthogonal antiparallel homodimers with Lys at the position to maintain the orthogonality within the set of the peptides. Lending the ideal of incorporation of a single charged residue at an interior position is sufficient to specify a dimeric structure.[69] To confirm that, the K22 residue at the fourth heptad the *a* position was mutated into an Ala (A) to decrease the electrostatic repulsion between the central heptads and tune the stability of the dimer. Therefore, two a-d' positions of APH4SN were mutated (K to A), and the resulting peptide, named APH4SN-A (Fig. 5b), was used as the new building block for the assembly of the designed TET12$_{2.3}$SN-f$_5$b cage, which was named TET12$_{2.3}$SN-f$_5$b-A.

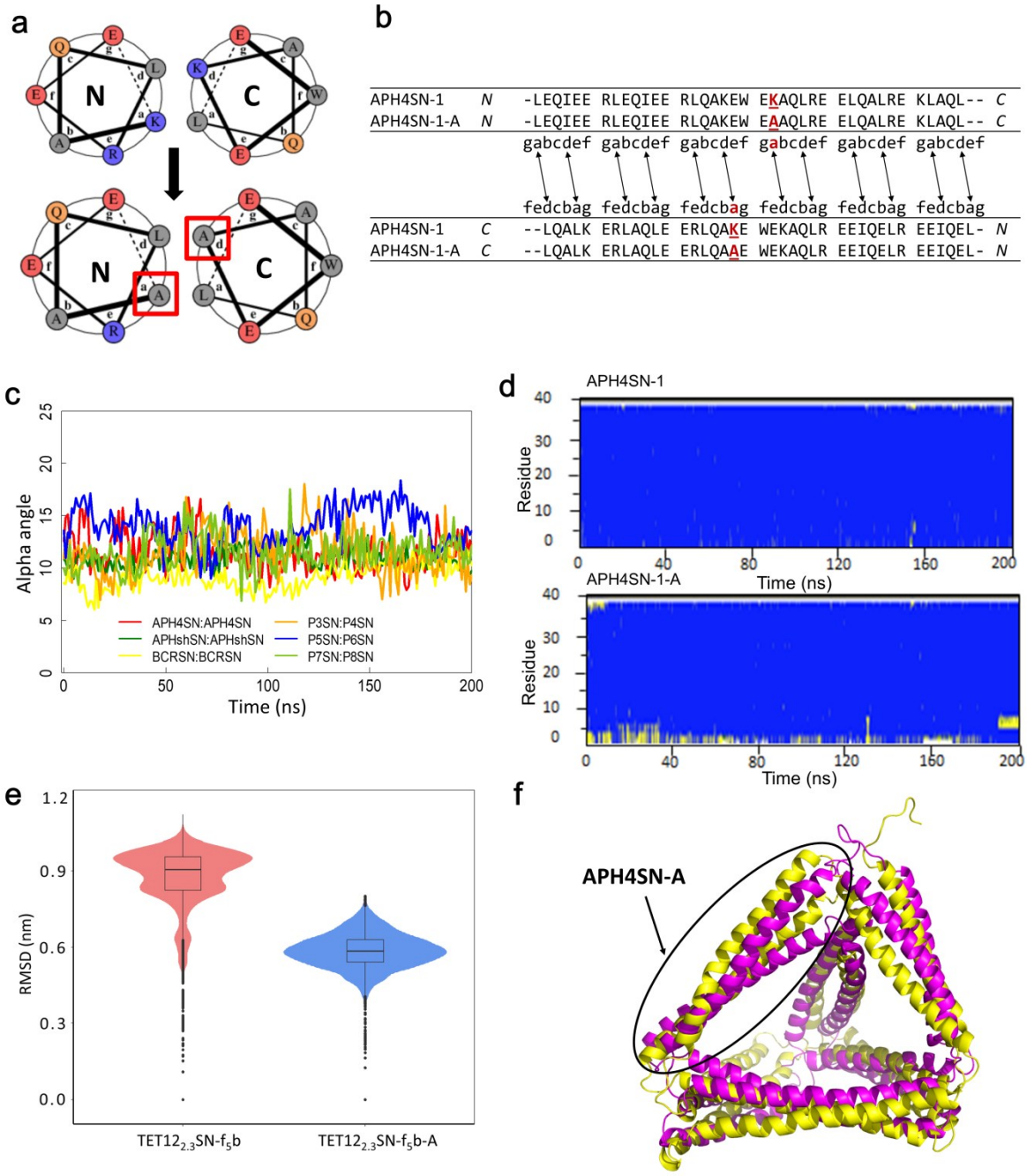


Figure 5. Rational redesign of APH4SN for conformational stability. (a) Helical wheel depiction of the coiled-coil dimer. (b) Aligned peptide sequences of APH4SN variants with K(A) positioned in a/d highlighted in a red box. (c) Alpha angle parameter for six coiled-coil segments of

TET12$_{2.3}$SN-f$_5$b-A, calculated over the simulation time. (d) Secondary structure assessment over time. (e) The RMSD distribution of TET12$_{2.3}$SN-f$_5$b and TET12$_{2.3}$SN-f$_5$b-A. (f) Superimposition of the initial (yellow) and final (magenta) structure of TET12$_{2.3}$SN-f$_5$b-A with cartoon representations. The APH4SN-A segment is highlighted by a black circle.

We also performed a 200ns MD simulation of the designed TET12$_{2.3}$SN-f$_5$b-A cage to investigate and compare its stability. To understand the dynamic effect of the molecular design of CCPO, we calculated the RMSFs (Fig. S9 (a)), structural parameters (Fig. 5c and Fig. S10), and second structure components (Fig. 5d), as well as the RMSDs (Fig. 5e), Rg, V, to describe the local and global stability of the original and designed CCPOs. In Fig S8, the designed APH4SN-A dimer shows RMSF values similar to those of other CC dimers, which means that the local stability of the APH4SN dimer is apparently strengthened due to the K-to-A mutation. The alpha-angles of the designed APH4SN-A dimer decrease (the red line in Fig. 5c), similar to other dimers in the designed TET12$_{2.3}$SN-f$_5$b-A cage. As shown in Fig. 5d, uncoiling deformations were not observed in either designed helical secondary structure of the designed APH4SN-A-1 and APH4SN-A-2 segments, and the higher values of the percentage of secondary structure of each segment in the designed cage are listed in Table S3. These local properties of the designed APH4SN dimer are reflected in the more regular packing model, and the helical secondary structure is well preserved over the simulations, providing confidence that this designed structure behaves as intended. Fig. 5e shows the RMSDs for TET12$_{2.3}$SN-f$_5$b cages with original and designed structures. The average RMSD values and their standard deviations during MD simulations for TET12$_{2.3}$SN-f$_5$b and TET12$_{2.3}$SN-f$_5$b-A are 0.88±0.11 nm and 0.58±0.07 nm, respectively. In addition, as shown in Fig. 5f and movie S4, the 3D structural mapping of the initial

and end structures at 200 ns means that the global structure of the designed TET12$_{2.3}$SN-f$_5$b cage is maintained during the MD simulation. As shown in Fig. S9 (b) and (c), TET12$_{2.3}$SN-f$_5$b-A shows similar gyration radius patterns but different volumes with TET12$_{2.3}$SN-f$_5$b. The distribution of volumes is smoother, which means a smaller deviation along the simulation. This finding also supports our design aim that the global structure of the designed TET12$_{2.3}$SN-f$_5$b cage indicates better structural stability due to residue modification from positively charged K to A.

4. Conclusions

In this work, we have investigated the stability and dynamic properties of 15 CCPO cages, in the shape of tetrahedra, square pyramids and triangular prisms, based on MD simulations. The good agreement between the SAXS experiment and MD simulation results indicates that most of the CCPO cages maintain their designed shapes during the simulation. One of the CC building blocks of the designed APH4SN segment was deformed in all cages that used this CC segment. By investigating the CC parameters and secondary structure component, the structural determinant and deformation mechanism of APH4sn are proposed. The redesigned CC dimer APH4SN-A (K to A) was used as a building block in the TET12$_{2.3}$SN-f$_5$b cage, and the simulation results show that both the CC dimer and cage structures are significantly improved and stabilized. In summary, we have proposed a molecular design strategy that relies on the mutation of the unstable CC dimer in the CCPO cages. Both the CC and secondary structure show that the mutation reduces conformational dynamics at the APH4SN interface and stabilizes the global conformation of the TET12$_{2.3}$SN-f$_5$b cage.

ASSOCIATED CONTENT

Supporting Information.

MD results for $TET12_{1.10}S\text{-}c_6$, $PYR16_{2.15R}S\text{-}f_5$ and $TRIP18_{7.5R}SN\text{-}f_5$-obl (Figure S1);

Time evolution of the RMSD calculated over the entire trajectory for all studied CCPO cages (Figure S2);

Time evolution of the Rg calculated over the entire trajectory for all studied CCPO cages (Figure S3);

Time evolution of the volume calculated over the entire trajectory for all studied CCPO cages (Figure S4);

Comparison of the MD simulation results and the experimental SAXS data (Figure S5);

The overlap map between the first PCA modes of $TET12_{1.10}SN\text{-}f_5^{best}$ and $TET12_{1.10}SN\text{-}f_5^{largest}$ (Figure S6);

RMSF calculated over the simulation time for each strand of all studied CCPO cages (Figure S7);

CC parameters for six coiled coil segments of $TET12_{2.3}SN\text{-}f_5b$-A (Figure S8);

MD results for $TET12_{2.3}SN\text{-}f_5b$-A (Figure S9);

CC parameters for six CC segments of $TET12_{2.3}SN\text{-}f_5b$-A (Figure S10);

RMSDs, Rg, and V between the initial and final conformations during MD simulations of each CCPO cage (Table S1);

Mean values of the eight parameters along the MD simulation with their standard deviations of six CC dimers (Table S2);

Secondary structure content of each segment in $TET12_{2.3}SN\text{-}f_5b$ and $TET12_{2.3}SN\text{-}f_5b$ -A (Table S3);

Dynamics of $TET12_{1.10}SN\text{-}f_5^{best}$ during the course of 200ns MD simulations (Movie S1);

Dynamics of $TET12_{1.10}SN\text{-}f_5^{largest}$ during the course of 200ns MD simulations (Movie S2);

Dynamics of original $TET12_{2.3}SN\text{-}f_5b$ during the course of 200ns MD simulation (Movie S3);

Dynamics of the designed TET12$_{2.3}$SN-f$_5$b during the course of 200ns MD simulation (Movie S4).

AUTHOR INFORMATION

Corresponding Author

Guang Hu - MOE Key Laboratory of Geriatric Diseases and Immunology, Suzhou Key Laboratory of Pathogen Bioscience and Anti-infective Medicine, Department of Bioinformatics, Center for Systems Biology, School of Biology and Basic Medical Sciences, Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Soochow University, Suzhou 215123, China;

Email: huguang@suda.edu.cn

Roman Jerala - Department of Synthetic Biology and Immunology, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia.

Email: roman.jerala@ki.si,

Authors

Fei Xiao - MOE Key Laboratory of Geriatric Diseases and Immunology, Suzhou Key Laboratory of Pathogen Bioscience and Anti-infective Medicine, Department of Bioinformatics, Center for Systems Biology, School of Biology and Basic Medical Sciences, Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Soochow University, Suzhou 215123, China

Longfei Luo - MOE Key Laboratory of Geriatric Diseases and Immunology, Suzhou Key Laboratory of Pathogen Bioscience and Anti-infective Medicine, Department of Bioinformatics,

Center for Systems Biology, School of Biology and Basic Medical Sciences, Suzhou Medical College of Soochow University, Suzhou, 215213, China

Xin Liu - Institute of Blood and Marrow Transplantation, Medical College of Soochow University, Jiangsu Institute of Hematology, The first Affiliated Hospital of Soochow University, Collaborative Innovation Center of Hematology, National Clinical Research Center for Hematologic Diseases, Soochow University, Suzhou 215123, China.

Ajasja Ljubetič - Department of Synthetic Biology and Immunology, National Institute of Chemistry, SI-1000 Ljubljana, Slovenia.

Nengzhi Jin - Key Laboratory of Advanced Computing of Gansu Province, Gansu Computing Center Lanzhou, 730030，China

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (32271292, 31872723), and a Project Funded by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions and Slovenian Research Agency project J1-4406.

ABBREVIATIONS

CCPO， coiled-coil protein origami； CC, coiled-coil; PCA, principal component analysis; SAXS, small angle X-ray scattering; RMSD, root mean square deviation; RMSF, root mean square fluctuation.

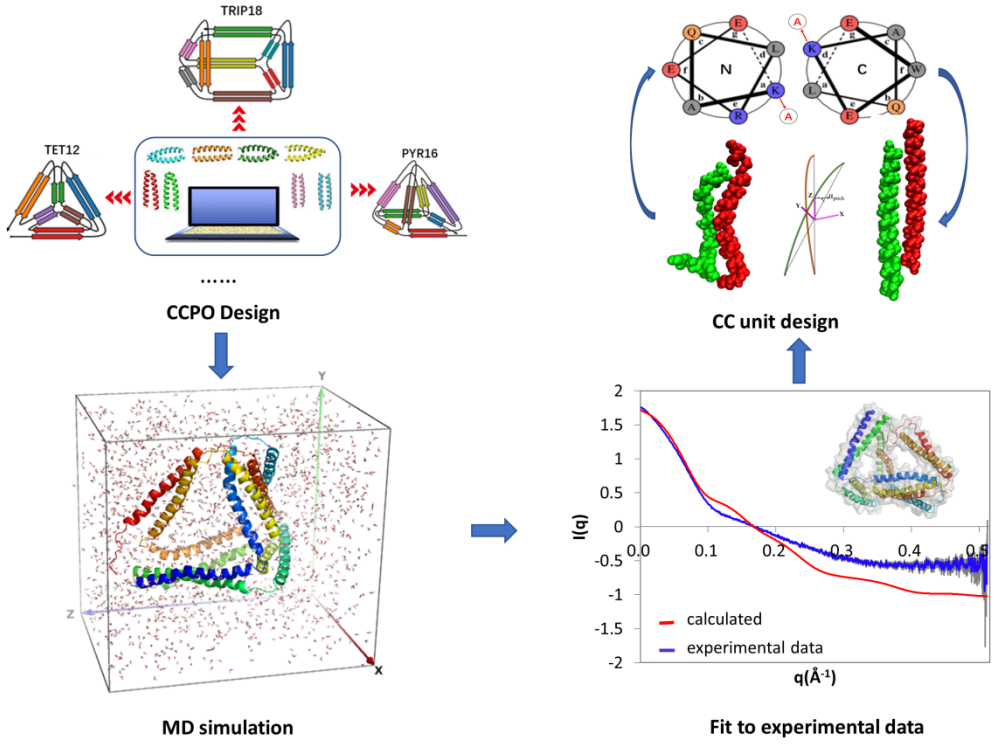
References

1. Jones, M. R.; Seeman, N. C.; Mirkin, C. A. Programmable Materials and the Nature of the DNA Bond. *SCIENCE* **2015,** *347*, No. 6224.
2. Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H. Structural DNA Nanotechnology: State of the Art and Future Perspective. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2014,** *136* (32), 11198-11211.
3. Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *SCIENCE* **2005,** *310* (5754), 1661-1665.
4. He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. Hierarchical Self-Assembly of DNA into Symmetric Supramolecular Polyhedra. *NATURE* **2008,** *452* (7184), 198-201.
5. Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. P.; et al. Self-Assembly of a Nanoscale DNA Box with a Controllable Lid. *NATURE* **2009,** *459* (7243), 73-75.
6. Gradisar, H.; Jerala, R. Self-Assembled Bionanostructures: Proteins Following the Lead of DNA Nanostructures. *JOURNAL OF NANOBIOTECHNOLOGY* **2014,** *12*, No. 4.
7. Yeates, T. O.; Liu, Y.; Laniado, J. The Design of Symmetric Protein Nanomaterials Comes of Age in Theory and Practice. *CURRENT OPINION IN STRUCTURAL BIOLOGY* **2016,** *39*, 134-143.
8. King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; Andre, I.; Gonen, T.; Yeates, T. O.; Baker, D. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *SCIENCE* **2012,** *336* (6085), 1171-1174.
9. Lai, Y.-T.; Cascio, D.; Yeates, T. O. Structure of a 16-Nm Cage Designed by Using Protein Oligomers. *SCIENCE* **2012,** *336* (6085), 1129-1129.
10. Lai, Y.-T.; Reading, E.; Hura, G. L.; Tsai, K.-L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O. Structure of a Designed Protein Cage That Self-Assembles into a Highly Porous Cube. *NATURE CHEMISTRY* **2014,** *6* (12), 1065-1071.
11. Lai, Y.-T.; Tsai, K.-L.; Sawaya, M. R.; Asturias, F. J.; Yeates, T. O. Structure and Flexibility of Nanoscale Protein Cages Designed by Symmetric Self-Assembly. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2013,** *135* (20), 7738-7743.
12. Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P.; et al. Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. *SCIENCE* **2016,** *353* (6297), 389-394.
13. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with Alphafold. *Nature* **2021,** *596* (7873), 583-589.
14. Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; et al. De Novo Design of Protein Structure and Function with Rfdiffusion. *Nature* **2023,** *620* (7976), 1089-1100.
15. Ingraham, J. B.; Baranov, M.; Costello, Z.; Barber, K. W.; Wang, W.; Ismail, A.; Frappier, V.; Lord, D. M.; Ng-Thow-Hing, C.; Van Vlack, E. R.; et al. Illuminating Protein Space with a Programmable Generative Model. *Nature* **2023,** *623* (7989), 1070-1078.
16. Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; et al. Robust Deep Learning-Based Protein Sequence Design Using Proteinmpnn. *SCIENCE* **2022,** *378* (6615), 49-55.

17. Lee, E. J.; Lee, N. K.; Kim, I.-S. Bioengineered Protein-Based Nanocage for Drug Delivery. *ADVANCED DRUG DELIVERY REVIEWS* **2016,** *106*, 157-171.
18. Vindigni, G.; Raniolo, S.; Ottaviani, A.; Falconi, M.; Franch, O.; Knudsen, B. R.; Desideri, A.; Biocca, S. Receptor-Mediated Entry of Pristine Octahedral DNA Nanocages in Mammalian Cells. *ACS NANO* **2016,** *10* (6), 5971-5979.
19. Zhao, Z.; Fu, J.; Dhakal, S.; Johnson-Buck, A.; Liu, M.; Zhang, T.; Woodbury, N. W.; Liu, Y.; Walter, N. G.; Yan, H. Nanocaged Enzymes with Enhanced Catalytic Activity and Increased Stability against Protease Digestion. *NATURE COMMUNICATIONS* **2016,** *7*, No. 10619.
20. Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *NATURE* **2006,** *440* (7082), 297-302.
21. Gradisar, H.; Bozic, S.; Doles, T.; Vengust, D.; Hafner-Bratkovic, I.; Mertelj, A.; Webb, B.; Sali, A.; Klavzar, S.; Jerala, R. Design of a Single-Chain Polypeptide Tetrahedron Assembled from Coiled-Coil Segments. *NATURE CHEMICAL BIOLOGY* **2013,** *9* (6), 362-366.
22. Kočar, V.; Abram, S. B.; Doles, T.; Bašić, N.; Gradišar, H.; Pisanski, T.; Jerala, R. Topofold, the Designed Modular Biomolecular Folds: Polypeptide-Based Molecular Origami Nanostructures Following the Footsteps of DNA. *WILEY INTERDISCIPLINARY REVIEWS-NANOMEDICINE AND NANOBIOTECHNOLOGY* **2015,** *7* (2), 218-237.
23. Grigoryan, G.; Keating, A. E. Structural Specificity in Coiled-Coil Interactions. *CURRENT OPINION IN STRUCTURAL BIOLOGY* **2008,** *18* (4), 477-483.
24. Ljubetic, A.; Drobnak, I.; Gradisar, H.; Jerala, R. Designing the Structure and Folding Pathway of Modular Topological Bionanostructures. *CHEMICAL COMMUNICATIONS* **2016,** *52* (30), 5220-5229.
25. Drobnak, I.; Ljubetic, A.; Gradisar, H.; Pisanski, T.; Jerala, R., Designed Protein Origami. In *Protein-Based Engineered Nanostructures*, Cortajarena, A. L.; Grove, T. Z., Eds. **2016**; *940*, 7-27.
26. Hu, G.; Qiu, W.-Y.; Ceulemans, A. A New Euler's Formula for DNA Polyhedra. *PLOS ONE* **2011,** *6*, No. 10.
27. Hu, G.; Wang, Z.; Qiu, W.-Y. A Survey on Mathematical Models for DNA Polyhedra. *MATCH-COMMUNICATIONS IN MATHEMATICAL AND IN COMPUTER CHEMISTRY* **2013,** *70* (3), 725-742.
28. Hu, G.; Zhai, X.-D.; Lu, D.; Qiu, W.-Y. The Architecture of Platonic Polyhedral Links. *JOURNAL OF MATHEMATICAL CHEMISTRY* **2009,** *46* (2), 592-603.
29. Fijavz, G.; Pisanski, T.; Rus, J. Strong Traces Model of Self-Assembly Polypeptide Structures. *MATCH-COMMUNICATIONS IN MATHEMATICAL AND IN COMPUTER CHEMISTRY* **2014,** *71* (1), 199-212.
30. Alves, C.; Iacovelli, F.; Falconi, M.; Cardamone, F.; della Rocca, B. M.; de Oliveira, C. L. P.; Desideri, A. A Simple and Fast Semiautomatic Procedure for the Atomistic Modeling of Complex DNA Polyhedra. *JOURNAL OF CHEMICAL INFORMATION AND MODELING* **2016,** *56* (5), 941-949.
31. Iacovelli, F.; Falconi, M. Decoding the Conformation-Linked Functional Properties of Nucleic Acids by the Use of Computational Tools. *FEBS JOURNAL* **2015,** *282* (17), 3298-3310.
32. Falconi, M.; Oteri, F.; Chillemi, G.; Andersen, F. F.; Tordrup, D.; Oliveira, C. L. P.; Pedersen, J. S.; Knudsen, B. R.; Desideri, A. Deciphering the Structural Properties That Confer Stability to a DNA Nanocage. *ACS NANO* **2009,** *3* (7), 1813-1822.
33. Iacovelli, F.; Alves, C.; Falconi, M.; Oteri, F.; de Oliveira, C. L. P.; Desideri, A. Influence of the Single-Strand Linker Composition on the Structural/Dynamical Properties of a Truncated Octahedral DNA Nano-Cage Family. *BIOPOLYMERS* **2014,** *101* (10), 992-999.
34. Iacovelli, F.; Falconi, M.; Knudsen, B. R.; Desideri, A. Comparative Simulative Analysis of Single and Double Stranded Truncated Octahedral DNA Nanocages. *RSC ADVANCES* **2016,** *6* (42), 35160-35166.
35. Oliveira, C. L. P.; Juul, S.; Jorgensen, H. L.; Knudsen, B.; Tordrup, D.; Oteri, F.; Falconi, M.; Koch, J.; Desideri, A.; Pedersen, J. S.; et al. Structure of Nanoscale Truncated Octahedral DNA Cages: Variation of Single-Stranded Linker Regions and Influence on Assembly Yields. *ACS NANO* **2010,** *4* (3), 1367-1376.
36. Oteri, F.; Falconi, M.; Chillemi, G.; Andersen, F. F.; Oliveira, C. L. P.; Pedersen, J. S.; Knudsen, B. R.; Desideri, A. Simulative Analysis of a Truncated Octahedral DNA Nanocage Family Indicates the Single-Stranded Thymidine Linkers as the Major Player for the Conformational Variability. *JOURNAL OF PHYSICAL CHEMISTRY C* **2011,** *115* (34), 16819-16827.

37. Franch, O.; Iacovelli, F.; Falconi, M.; Juul, S.; Ottaviani, A.; Benvenuti, C.; Biocca, S.; Ho, Y.-P.; Knudsen, B. R.; Desideri, A. DNA Hairpins Promote Temperature Controlled Cargo Encapsulation in a Truncated Octahedral Nanocage Structure Family. *NANOSCALE* **2016,** *8* (27), 13333-13341.
38. Juul, S.; Iacovelli, F.; Falconi, M.; Kragh, S. L.; Christensen, B.; Frohlich, R.; Franch, O.; Kristoffersen, E. L.; Stougaard, M.; Leong, K. W.; et al. Temperature-Controlled Encapsulation and Release of an Active Enzyme in the Cavity of a Self-Assembled DNA Nanocage. *ACS NANO* **2013,** *7* (11), 9724-9734.
39. Yoo, J.; Aksimentiev, A. In Situ Structure and Dynamics of DNA Origami Determined through Molecular Dynamics Simulations. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA* **2013,** *110* (50), 20099-20104.
40. Doye, J. P. K.; Ouldridge, T. E.; Louis, A. A.; Romano, F.; Sulc, P.; Matek, C.; Snodin, B. E. K.; Rovigatti, L.; Schreck, J. S.; Harrison, R. M.; et al. Coarse-Graining DNA for Simulations of DNA Nanotechnology. *PHYSICAL CHEMISTRY CHEMICAL PHYSICS* **2013,** *15* (47), 20395-20414.
41. Schreck, J. S.; Romano, F.; Zimmer, M. H.; Louis, A. A.; Doye, J. P. K. Characterizing DNA Star-Tile-Based Nanostructures Using a Coarse-Grained Model. *ACS NANO* **2016,** *10* (4), 4236-4247.
42. Snodin, B. E. K.; Romano, F.; Rovigatti, L.; Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. Direct Simulation of the Self-Assembly of a Small DNA Origami. *ACS NANO* **2016,** *10* (2), 1724-1737.
43. Kočar, V.; Schreck, J. S.; Čeru, S.; Gradišar, H.; Bašić, N.; Pisanski, T.; Doye, J. P. K.; Jerala, R. Design Principles for Rapid Folding of Knotted DNA Nanostructures. *Nature Communications* **2016,** *7* (1).
44. Satler, T.; Hadzi, S.; Jerala, R. Crystal Structure of *De Novo* Designed Coiled-Coil Protein Origami Triangle. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2023,** *145* (31), 16995-17000.
45. Ljubetic, A.; Lapenta, F.; Gradisar, H.; Drobnak, I.; Aupic, J.; Strmsek, Z.; Lainscek, D.; Hafner-Bratkovic, I.; Majerle, A.; Krivec, N.; et al. Design of Coiled-Coil Protein-Origami Cages That Self-Assemble in Vitro and in Vivo. *NATURE BIOTECHNOLOGY* **2017,** *35* (11), 1094-1101.
46. Kutzner, C.; Pall, S.; Fechner, M.; Esztermann, A.; de Groot, B. L.; Grubmueller, H. Best Bang for Your Buck: Gpu Nodes for Gromacs Biomolecular Simulations. *JOURNAL OF COMPUTATIONAL CHEMISTRY* **2015,** *36* (26), 1990-2008.
47. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99sb Protein Force Field. *PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS* **2010,** *78* (8), 1950-1958.
48. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *JOURNAL OF CHEMICAL PHYSICS* **1993,** *98* (12), 10089-10092.
49. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. Amber, a Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *COMPUTER PHYSICS COMMUNICATIONS* **1995,** *91* (1-3), 1-41.
50. Andersen, H. C. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *JOURNAL OF COMPUTATIONAL PHYSICS* **1983,** *52* (1), 24-34.
51. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *JOURNAL OF COMPUTATIONAL PHYSICS* **1977,** *23* (3), 327-341.
52. Le Grand, S.; Goetz, A. W.; Walker, R. C. Spfp: Speed without Compromise-a Mixed Precision Model for Gpu Accelerated Molecular Dynamics Simulations. *COMPUTER PHYSICS COMMUNICATIONS* **2013,** *184* (2), 374-380.
53. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *WILEY INTERDISCIPLINARY REVIEWS-COMPUTATIONAL MOLECULAR SCIENCE* **2013,** *3* (2), 198-210.
54. Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *JOURNAL OF MOLECULAR GRAPHICS & MODELLING* **1996,** *14* (1), 33-38.
55. DeLano, W. L.; Lam, J. W. Pymol: A Communications Tool for Computational Models. *ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY* **2005,** *230*, U1371-U1372.

56. Karplus, M.; Ichiye, T. Fluctuation and Cross Correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations. *JOURNAL OF MOLECULAR BIOLOGY* **1996,** *263* (2), 120-122.
57. Raghavendra Sridharamurthy; Talha Bin Masood; Harish Doraiswamy; Siddharth Patel; Raghavan Varadarajan; Natarajan, V., Extraction of Robust Voids and Pockets in Proteins. In *Visualization in Medicine and Life Sciences Iii*, 1 ed.; Linsen, L.; Hamann, B.; Hege, H.-C., Eds. Springer International Publishing: **2016**, VIII, 352.
58. Svergun, D.; Barberato, C.; Koch, M. H. J. Crysol - a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *JOURNAL OF APPLIED CRYSTALLOGRAPHY* **1995,** *28*, 768-773.
59. Crick, F. H. C. The Fourier Transform of a Coiled-Coil. *ACTA CRYSTALLOGRAPHICA* **1953,** *6* (8-9), 685-689.
60. Offer, G.; Hicks, M. R.; Woolfson, D. N. Generalized Crick Equations for Modeling Noncanonical Coiled Coils. *JOURNAL OF STRUCTURAL BIOLOGY* **2002,** *137* (1-2), 41-53.
61. Grigoryan, G.; DeGrado, W. F. Probing Designability Via a Generalized Model of Helical Bundle Geometry. *JOURNAL OF MOLECULAR BIOLOGY* **2011,** *405* (4), 1079-1100.
62. Wood, C. W.; Bruning, M.; Ibarra, A. A.; Bartlett, G. J.; Thomson, A. R.; Sessions, R. B.; Brady, R. L.; Woolfson, D. N. Ccbuilder: An Interactive Web-Based Tool for Building, Designing and Assessing Coiled-Coil Protein Assemblies. *BIOINFORMATICS* **2014,** *30* (21), 3029-3035.
63. Wood, C. W.; Heal, J. W.; Thomson, A. R.; Bartlett, G. J.; Ibarra, A. A.; Brady, R. L.; Sessions, R. B.; Woolfson, D. N. Isambard: An Open-Source Computational Environment for Biomolecular Analysis, Modelling and Design. *BIOINFORMATICS* **2017,** *33* (19), 3043-3050.
64. Walshaw, J.; Woolfson, D. N. Socket: A Program for Identifying and Analysing Coiled-Coil Motifs within Protein Structures. *JOURNAL OF MOLECULAR BIOLOGY* **2001,** *307* (5), 1427-1450.
65. Lapenta, F.; Aupic, J.; Strmsek, Z.; Jerala, R. Coiled Coil Protein Origami: From Modular Design Principles Towards Biotechnological Applications. *CHEMICAL SOCIETY REVIEWS* **2018,** *47* (10), 3530-3542.
66. Boldridge, W. C.; Ljubetič, A.; Kim, H.; Lubock, N.; Szilágyi, D.; Lee, J.; Brodnik, A.; Jerala, R.; Kosuri, S. A Multiplexed Bacterial Two-Hybrid for Rapid Characterization of Protein–Protein Interactions and Iterative Protein Design. *Nature Communications* **2023,** *14* (1).
67. Plaper, T.; Aupič, J.; Dekleva, P.; Lapenta, F.; Keber, M. M.; Jerala, R.; Benčina, M. Coiled-Coil Heterodimers with Increased Stability for Cellular Regulation and Sensing Sars-Cov-2 Spike Protein-Mediated Cell Fusion. *Scientific Reports* **2021,** *11* (1), No. 9136.
68. Drobnak, I.; Gradisar, H.; Ljubetic, A.; Merljak, E.; Jerala, R. Modulation of Coiled-Coil Dimer Stability through Surface Residues While Preserving Pairing Specificity. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2017,** *139* (24), 8229-8236.
69. Negron, C.; Keating, A. E. A Set of Computationally Designed Orthogonal Antiparallel Homodimers That Expands the Synthetic Coiled-Coil Toolkit. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2014,** *136* (47), 16544-16556.


TRIP18
TET12
PYR16
.....
CCPO Design
MD simulation
CC unit design
Fit to experimental data
N
C
I(q)
q(Å-1)
calculated
experimental data

TOC Graphic