

AlphaFold 3 – Aided Design of DNA Motifs To Assemble into Triangles

Anusha, Zhe Zhang, Jinyue Li, Hua Zuo,* and Chengde Mao*

Cite This: *J. Am. Chem. Soc.* 2024, 146, 25422–25425

Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: Self-assembly of biomolecules provides a powerful tool for a wide range of applications in nanomedicine, biosensing and imaging, vaccines, computation, nanophotonics, etc. The key is to rationally design building blocks and the intermolecule interactions. Along this line, structural DNA nanotechnology has rapidly developed by limiting DNA secondary structures to primarily well-established, B-form DNA duplexes, which can be readily and reliably predicted. As the field evolves, more sophisticated structural elements must be introduced. While increasing the structural complexity, they bring challenges to predicting DNA nanostructures. In the past, a brutal and tedious error-and-trial approach has often been used to solve this problem. Here, we report a case study of applying AlphaFold 3 to model the structural elements to facilitate DNA nanostructure design. This protocol is expected to be generally applicable and greatly facilitates the further development of structural DNA nanotechnology.

Programmed self-assembly of biomolecules, e.g. DNA, provided a powerful approach to prepare structures across multiple size scales.^{1–15} The foundation of this approach is the knowledge accumulated in structural biology,¹⁶ and the challenges are often due to the lack of detailed structural information. Limited by our knowledge, the physical realization of the designs often requires a brutal and tedious trial-and-error approach to examine multiple choices of each design parameter. With the recent release of AlphaFold 3 (AF3),¹⁷ an integrated structural prediction algorithm, we should be able to model the designs in silico to quickly find the solution, thus, greatly reducing the experimental works and speeding up the research process. Herein, we showcase a study that shows that AF3 helps to optimize the parameters of a DNA nanomotif to allow the nanomotifs to self-assemble into the designed DNA nanotriangles that can be effectively ligated.

A triangle, in principle, could be self-assembled from a DNA motif with a 60°-bend (Figure 1). The motifs can form triangles only if (1) each motif bends 60°, (2) the loops bend toward the same direction, and (3) the motifs are on the same

plane. The simplest motif of this type is a DNA duplex with a single-stranded loop in the center. The loop forces the DNA duplex to bend toward the face of the DNA duplex opposite from the loop. By mimicking an RNA k-turn motif,¹⁸ we have identified a 5-nucleotide (nt)-long loop, AACTA, to bend the DNA duplex roughly of 60°. When such motifs associate with each other, if the loops are separated by integral numbers of helical turns, the bends will accumulate in the same direction. If being on the same plane, the associated motifs will form a closed complex (triangle); otherwise, they will form open-ended, spirals.¹⁸

We have used Tiamat,²⁰ a DNA modeling software, to design a DNA nanomotif to assemble a DNA triangle with edges of two turns long. The edge length is chosen for the ease of DNA synthesis and the sufficient motif stability. The motif contains a red strand Rx and a blue strand Bx. x indicates the number of base pairs of each edge. Initially, we assigned $x = 21$ base pairs (bps) for two helical turns;²¹ thus, the two strands, R21 and B21, are 21 and $21 + 5 = 26$ nts long, respectively. A 5-nt-loop is located at the center of the motif, and a pair of complementary sticky ends are at the two ends of the motif.

The DNA complexes were prepared by thermally annealing from 95° to 24 °C and examined by native polyacrylamide gel electrophoresis (nPAGE) at 24 °C (Figure 2f). To facilitate interpreting the nPAGE data, three control motifs were prepared: (i) a motif with one blunt end (B21* + R21). It could form the motif but could not oligomerize, thus existing as a monomer motif. (ii) a twice motif (B19 + R19x2 at molar

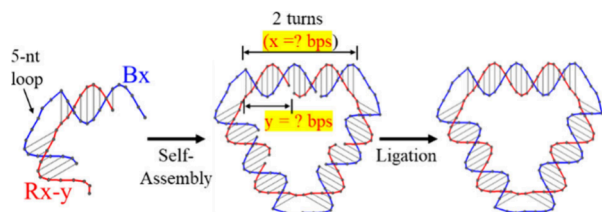


Figure 1. Scheme of assembly of a DNA triangle from a bulged duplex nanomotif. It is a two-turn-long DNA duplex containing a 5-nucleotide (nt)-long loop at the center and a pair of complementary sticky ends at two ends. x: the number of base pairs (bps) of each edge; y: the number of bps between the 5' end of the red strand and the bulge loop. The triangle can be further stabilized by enzymatic ligation with the T4 DNA ligase.

Received: June 20, 2024

Revised: August 27, 2024

Accepted: August 28, 2024

Published: September 5, 2024



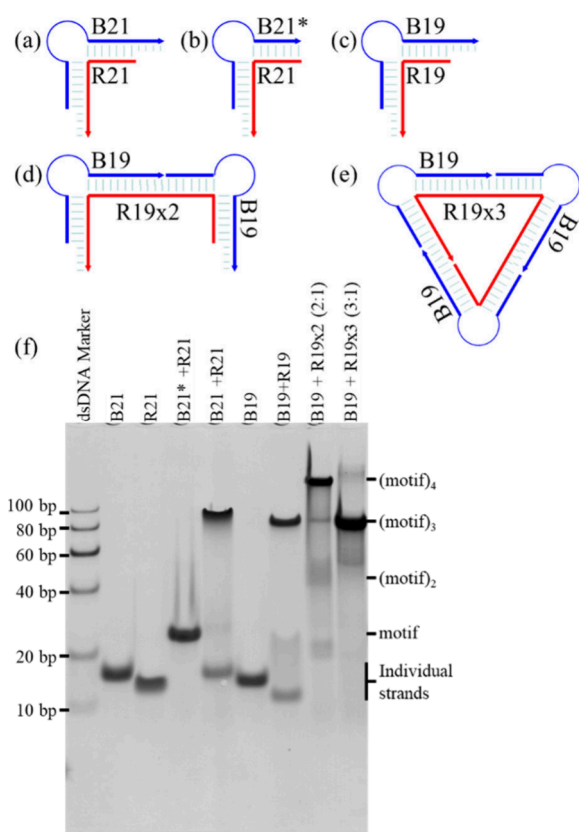


Figure 2. Analysis of the DNA triangle assembly by 10% native polyacrylamide gel electrophoresis (nPAGE). (a–e) DNA motifs. (f) nPAGE image. The sample compositions and chemical identities of the bands are indicated above and beside the gel image, respectively.

ratio of 2:1). The red strand (R19x2) was 38 nts long, nearly twice as long as R21 (21 nts). Depending on association within one or between two motifs, the resulting complex will have similar shape and molecular weight to a dimer, (motif)₂, or a tetramer, (motif)₄, of the R21–B21 motif, respectively. And (iii) a thrice motif (B19 + R19x3 at a molar ratio of 3:1). The red strand was 57 nts long, nearly three times longer than R21 (21 nts). The three motifs have similar shape and molecular weight to a trimer of the R21–B21 motif, (motif)₃. From the nPAGE data (Figure 2f), we could see that the DNA sample (R21 + B21) indeed formed a major band corresponding to the designed triangle, (motif)₃. However, the band is not sharp, and a continuous smear shows up below the major band, suggesting that the DNA triangle was not stable and constantly dissociated into small fragments to form the smear in the gel electrophoresis.

To understand why the triangle is not stable, we used AF3 to model the association among motifs (Figure 3). The main uncertainty is how the loop will influence the motif structure. We first modeled the structure of a DNA complex of two associated motifs. A simplified complex is used to mimic two associated motifs. It contains three duplex domains and two 5-nt-long loops. The central helical domain is 21 bps long. From the modeled structure, clearly the complex is not flat. The dihedral angle between the left and right helical domains is $\sim 64^\circ$. Thus, the 5-nt loop not only bends the duplex but also introduces an $\sim 64^\circ$ twist around the helix (equivalent to the helical twist of 2 bps, $360^\circ \times 2/10.5 = 68.6^\circ$).²² For three motifs to assemble into a triangle, each motif needs to

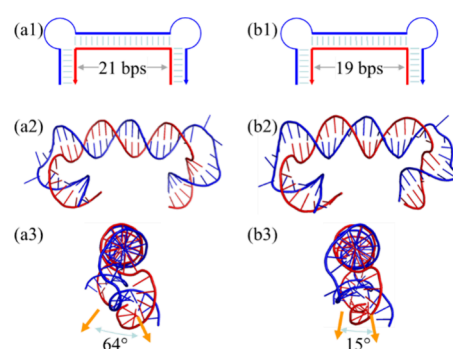


Figure 3. Modeling the structures of two associated, bulged DNA duplex motifs with AF3. The edge length $x = 21$ base pairs, bps (a) or $x = 19$ (b). Top panel (a1) and (b1): show the secondary structures of the DNA complexes. Middle panel, (a2) and (b2), and bottom, (a3) and (b3), show the AF3-predicted structures viewed perpendicular to or along the central duplex domains, respectively.

overcome the 64° overtwist and forces all DNA helical domains to sit on the same plane. This costs extra energy and makes the triangle unstable. Based on the modeling data, we reduced the length of the central helical domain by 2 bps from 21 to 19 bps. The revised complex was then modeled by AF3. In the resulting model, the dihedral angle between the left and right helical domains is $\sim 15^\circ$. The entire complex is almost flat.

Based on the AF3 results, we have revised the motif to a 19-bp-long edge (Figure 2c) of R19 and B19, which were 19 and 24 nts long, respectively. nPAGE (Figure 2f) analysis showed that R19 and B19 indeed assembled into the designed triangle, which was stable and appeared as a very sharp, strong band. AFM imaging further confirmed the triangle geometry (Figure S1).

T4 DNA ligase-mediated ligation is an efficient means to stabilize self-assembled DNA nanostructures.^{23–27} For ligase to work efficiently, it is important to make sure that ligase has access to the ligation site of the DNA substrate.²⁶ To find the proper ligation sites on the DNA triangle, we generated structural models of a series of DNA triangles with different nick positions (potential ligation sites). The distance between the 5' end of strand R19 and the bulge is y bps (thus, the corresponding red strand is dubbed R19- y). The nick site is a few bps away from the bulge. In the literature, a crystal structure (pdb id: 6dt1) has been reported for DNA T4 ligase in complex with a nicked DNA duplex.²⁸ We have aligned the crystal structure of T4 DNA ligase to the DNA triangle structure by assuming different ligation sites along the triangle edge (Figure 4). When $y = 6$, T4 DNA ligase has substantial overlap with the DNA triangle (Figure 4c2); thus, steric hindrance would prevent T4 DNA ligase from accessing the nick site for ligation. Consequently, a low ligation efficiency is expected. In contrast, when $y = 7$, T4 DNA ligase can readily access the nick position (Figure 4d); thus, a high ligation yield is expected. When comparing this series of aligned structural models, the access of T4 DNA ligase to the triangles decreases as $y = 7 > 8 \sim 9 > 6 > 10 \sim 11$. The ligation efficiency is expected to follow the same trend.

We experimentally confirmed the modeling prediction (Figure 5). All R19 variants were phosphorylated and annealed with equal molar concentrations of strand B19 separately. Then the ligation was conducted by adding T4 DNA ligase and being incubated at 24°C for 16 h. Finally, the ligation mixtures

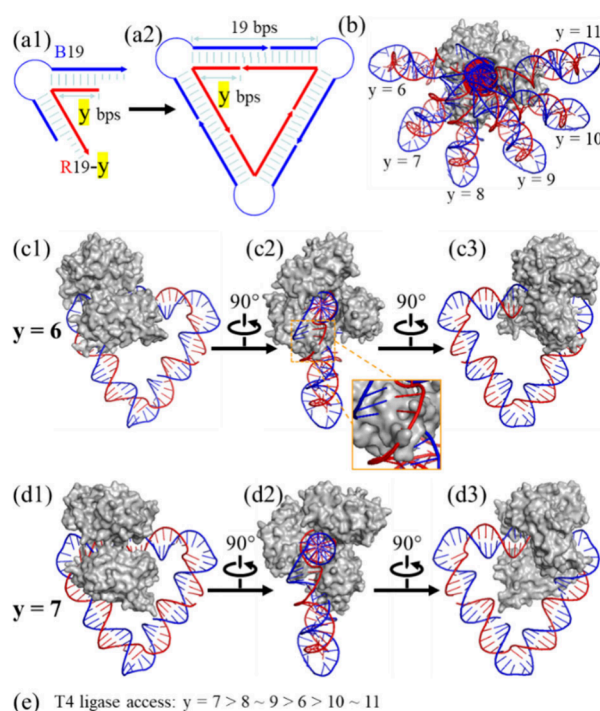


Figure 4. Structural alignment between the crystal structure of a T4 DNA ligase (pdb id: 6dt1) and AF3-predicted DNA triangles with nicks (potential ligation sites). (a) DNA motifs assemble into a two-turn-long triangle. Note that the S' end of strand R19-y is y bps away from the loop. (b) Superimposed T4-DNA ligase and six DNA triangles with different y values. T4 DNA ligase structure aligned with triangles with y values of (c) 6 and (d) 7. Note the substantial space clash in (c2).

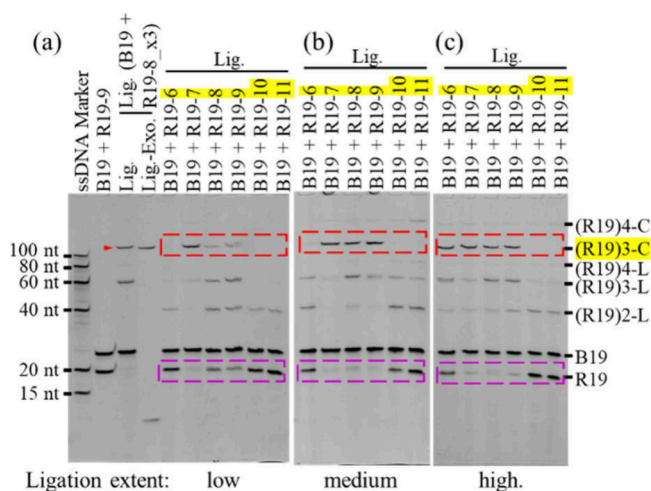


Figure 5. Enzymatic ligation of DNA triangles analyzed by denaturing PAGE (15%). From (a) to (b) to (c): ligation extent increases. The red and purple boxes indicate the starting DNA strands and the ligation products, respectively. The sample compositions and chemical identities of the bands are indicated above and beside the gel image, respectively.

were analyzed by denaturing PAGE. Upon ligation, the bands of the starting DNA strands (indicated by purple boxes) became weak, and the bands of the ligation products, circular R19 trimer [(R19)3-C] (indicated by red boxes) appeared. Consistent with what modeling suggested, the ligation efficiency was the highest when $y = 7$ and the lowest when y

$= 10$ or 11 . For the nonoptimal nick positions (e.g., $y = 6, 8$, and 9), ligation required an increased amount of T4 DNA ligase. An nPAGE was used to confirm that all the triangle variants could be assembled and were stable (Figure S2), thus, excluding the possibilities that ligation efficiency variation is due to the stabilities of the triangles.

Our preliminary study shows that the single-stranded loop (particularly the length) in the motif will influence the motif conformation (Figure S3). We have also demonstrated that, with a small revision, the basic bulge duplex motif can be designed to assemble into other nanostructures, e.g., a tetrameric square (Figure S4).

We have successfully applied AF3 to aid the design of DNA motifs for nanostructure assembly. Aside from the success in construction of the reported specific DNA nanostructures, this work has demonstrated the potential of AF3 to facilitate the design of DNA nanostructures. Structural modeling has long been used for DNA nanostructures from facilitating designs to predicting mechanical properties.^{29,30} In general, those softwares require certain prestudying and are not particularly easy to implement for tile-based self-assembly of DNA nanostructures by people with minimal computer skills. In contrast, AF3 is readily accessible, experimentalist-friendly, and easy to use. Moreover, it provides an integrated system to model complex systems of DNA, RNA, proteins, and ligands.¹⁷ It is conceivable that the reported design protocol can be applied to general and complex biomolecular self-assemblies.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.4c08387>.

Materials and detailed experimental methods; and four figures (AFM images and electrophoresis data) for additional characterization of DNA assembly. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Hua Zuo – College of Pharmaceutical Sciences, Southwest University, Chongqing 400715, China; orcid.org/0000-0002-8461-8988; Email: zuohua@swu.edu.cn

Chengde Mao – Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States; orcid.org/0000-0001-7516-8666; Email: mao@purdue.edu

Authors

Anusha – Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

Zhe Zhang – College of Pharmaceutical Sciences, Southwest University, Chongqing 400715, China; orcid.org/0000-0001-7465-163X

Jinyue Li – Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/jacs.4c08387>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by NSF (CCF-2107393 and CCF-2025187) and NSFC (22134005 and 22274315).

REFERENCES

- (1) Seeman, N. C.; Sleiman, H. F. DNA nanotechnology. *Nat. Rev. Mater.* **2018**, *3*, 17068.
- (2) Jones, M. R.; Seeman, N. C.; Mirkin, C. A. Programmable Materials and the Nature of the DNA Bond. *Science* **2015**, *347*, 840–840.
- (3) Hong, F.; Zhang, F.; Liu, Y.; Yan, H. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem. Rev.* **2017**, *117*, 12584–12640.
- (4) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. Challenges and Opportunities for Structural DNA Nanotechnology. *Nat. Nanotechnol.* **2011**, *6*, 763–772.
- (5) Ramezani, H.; Dietz, H. Building Machines with DNA Molecules. *Nat. Rev. Genet.* **2020**, *21*, 5–26.
- (6) Dong, Y.; Yao, C.; Zhu, Y.; Yang, L.; Luo, D.; Yang, D. DNA Functional Materials Assembled from Branched DNA: Design, Synthesis, and Applications. *Chem. Rev.* **2020**, *120*, 9420–9481.
- (7) Rothmund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297–302.
- (8) Ong, L. L.; Hanikel, N.; Yaghi, O. K.; Grun, C.; Strauss, M. T.; Bron, P.; Lai-Kee-Him, J.; Schueder, F.; Wang, B.; Wang, P.; Kishi, J. Y.; Myhrvold, C. A.; Zhu, A.; Jungmann, R.; Bellot, G.; Ke, Y.; Yin, P. Programmable Self-Assembly of Three-Dimensional Nanostructures from 10,000 Unique Components. *Nature* **2017**, *552*, 72–77.
- (9) Tikhomirov, G.; Petersen, P.; Qian, L. Fractal Assembly of Micrometre-Scale DNA Origami Arrays with Arbitrary Patterns. *Nature* **2017**, *552*, 67–71.
- (10) Pumm, A.-K.; Engelen, W.; Kopperger, E.; Isensee, J.; Vogt, M.; Kozina, V.; Kube, M.; Honemann, M. N.; Bertolin, E.; Langecker, M.; Golestanian, R.; Simmel, F. C.; Dietz, H. A DNA Origami Rotary Ratchet Motor. *Nature* **2022**, *607*, 492–498.
- (11) Li, Z.; Wang, S.; Nattermann, U.; Bera, A. K.; Borst, A. J.; Yaman, M. Y.; Bick, M. J.; Yang, E. C.; Sheffler, W.; Lee, B.; Seifert, S.; Hura, G. L.; Nguyen, H.; Kang, A.; Dalal, R.; Lubner, J. M.; Hsia, Y.; Haddox, H.; Courbet, A.; Dowling, Q.; Miranda, M.; Favor, A.; Etemadi, A.; Edman, N. I.; Yang, W.; Weidle, C.; Sankaran, B.; Negahdari, B.; Ross, M. B.; Ginger, D. S.; Baker, D. Accurate Computational Design of Three-Dimensional Protein Crystals. *Nat. Mater.* **2023**, *22*, 1556–1563.
- (12) Posnjak, G.; Yin, X.; Butler, P.; Bienek, O.; Dass, M.; Lee, S.; Sharp, I. D.; Liedl, T. Diamond-Lattice Photonic Crystals Assembled from DNA Origami. *Science* **2024**, *384*, 781–785.
- (13) Liu, H.; Matthies, M.; Russo, J.; Rovigatti, L.; Narayanan, R. P.; Diep, T.; McKeen, D.; Gang, O.; Stephanopoulos, N.; Sciortino, F.; Yan, H.; Romano, F.; Sulc, P. Inverse Design of a Pyrochlore Lattice of DNA Origami through Model-Driven Experiments. *Science* **2024**, *384*, 776–781.
- (14) Zhang, T.; Qian, X.; Zeng, W.; Wei, B. Custom Folding of Double-Stranded DNA Directed by Triplex Formation. *Chem* **2023**, *9*, 1505–1517.
- (15) Ng, C.; Samanta, A.; Mandrup, O. A.; Tsang, E.; Youssef, S.; Klausen, L. H.; Dong, M.; Nijenhuis, M. A. D.; Gothelf, K. V. Folding Double-Stranded DNA into Designed Shapes with Triplex-Forming Oligonucleotides. *Adv. Mater.* **2023**, *35*, 2302497.
- (16) Zardecki, C.; Dutta, S.; Goodsell, D. S.; Lowe, R.; Voigt, M.; Burley, S. K. PDB-101: Educational Resources Supporting Molecular Explorations through Biology and Medicine. *Protein Sci.* **2022**, *31*, 129–140.
- (17) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Zemgulyte, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlín, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Zidek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630*, 493.
- (18) Dibrov, S. M.; McLean, J.; Parsons, J.; Hermann, T. Self-Assembling RNA Square. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6405–6408.
- (19) Zuo, H.; Wu, S.; Li, M.; Li, Y.; Jiang, W.; Mao, C. A Case Study of the Likes and Dislikes of DNA and RNA in Self-Assembly. *Angew. Chem., Int. Ed.* **2015**, *54*, 15118–15121.
- (20) Williams, S.; Lund, K.; Lin, C.; Wonka, P.; Lindsay, S.; Yan, H. Tiamat: A Three-Dimensional Editing Tool for Complex DNA Structures. In *DNA Computing. DNA 2008, Lecture Notes in Computer Science*; Goel, A., Simmel, F. C., Sosik, P., Eds.; Springer: Berlin, 2009; pp 90–101.
- (21) Wang, J. C. Helical Repeat of DNA in Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 200–203.
- (22) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
- (23) Gerling, T.; Kube, M.; Kick, B.; Dietz, H. Sequence-Programmable Covalent Bonding of Designed DNA Assemblies. *Sci. Adv.* **2018**, *4*, No. eaau1157.
- (24) Orun, A. R.; Dmytriw, S.; Vajapayajula, A.; Snow, C. D. Stabilizing DNA-Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals* **2022**, *12*, 49.
- (25) Li, T.; Liu, D.; Chen, J.; Lee, A. H. F.; Qi, J.; Chan, A. S. C. Construction of Circular Oligodeoxynucleotides on the New Structural Basis of i-Motif. *J. Am. Chem. Soc.* **2001**, *123*, 12901–12902.
- (26) Li, Z.; Liu, L.; Zheng, M.; Zhao, J.; Seeman, N. C.; Mao, C. Making Engineered 3D DNA Crystals Robust. *J. Am. Chem. Soc.* **2019**, *141*, 15850–15855.
- (27) O'Neill, P.; Rothmund, P. W. K.; Kumar, A.; Fygenson, D. K. Sturdier DNA Nanotubes via Ligation. *Nano Lett.* **2006**, *6*, 1379–1383.
- (28) Shi, K.; Bohl, T. E.; Park, J.; Zasada, A.; Malik, S.; Banerjee, S.; Tran, V.; Li, N.; Yin, Z.; Kurniawan, F.; Orellana, K.; Aihara, H. T4 DNA Ligase Structure Reveals a Prototypical ATP-Dependent Ligase with a Unique Mode of Sliding Clamp Interaction. *Nucleic Acids Res.* **2018**, *46*, 10474–10488.
- (29) Poppleton, E.; Bohlin, J.; Matthies, M.; Sharma, S.; Zhang, F.; Sulc, P. Design, Optimization, and Analysis of Large DNA and RNA Nanostructures through Interactive Visualization, Editing, and Molecular Simulation. *Nucleic Acids Res.* **2020**, *48*, No. e72.
- (30) Douglas, S. M.; Marblestone, A. H.; Teerapittayanon, S.; Vazquez, A.; Church, G. M.; Shih, W. M. Rapid Prototyping of 3D DNA-Origami Shapes with caDNAno. *Nucleic Acids Res.* **2009**, *37*, 5001–5006.