






# The influence of Holliday junction sequence and dynamics on DNA crystal self-assembly

Chad R. Simmons<sup>1,5</sup>, Tara MacCulloch<sup>1,2,5</sup>, Miroslav Krepl<sup>3,4</sup> , Michael Matthies<sup>1</sup>, Alex Buchberger<sup>1,2</sup>, Ilyssa Crawford<sup>1,2</sup>, Jiří Šponer<sup>3,4</sup> , Petr Šulc<sup>1,2</sup> , Nicholas Stephanopoulos<sup>1,2</sup>  <sup>✉</sup> & Hao Yan<sup>1,2</sup>  <sup>✉</sup>

The programmable synthesis of rationally engineered crystal architectures for the precise arrangement of molecular species is a foundational goal in nanotechnology, and DNA has become one of the most prominent molecules for the construction of these materials. In particular, branched DNA junctions have been used as the central building block for the assembly of 3D lattices. Here, crystallography is used to probe the effect of all 36 immobile Holliday junction sequences on self-assembling DNA crystals. Contrary to the established paradigm in the field, most junctions yield crystals, with some enhancing the resolution or resulting in unique crystal symmetries. Unexpectedly, even the sequence adjacent to the junction has a significant effect on the crystal assemblies. Six of the immobile junction sequences are completely resistant to crystallization and thus deemed “fatal,” and molecular dynamics simulations reveal that these junctions invariably lack two discrete ion binding sites that are pivotal for crystal formation. The structures and dynamics detailed here could be used to inform future designs of both crystals and DNA nanostructures more broadly, and have potential implications for the molecular engineering of applied nanoelectronics, nanophotonics, and catalysis within the crystalline context.

<sup>1</sup>Biodesign Center for Molecular Design and Biomimetics, Arizona State University, 1001S. McAllister Ave, Tempe, AZ 85287, USA. <sup>2</sup>School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA. <sup>3</sup>Institute of Biophysics of the Czech Academy of Sciences, Královopolská 135, 612 65, Brno, Czech Republic. <sup>4</sup>Regional Centre of Advanced Technologies and Materials, Czech Advanced Technology and Research Institute (CATRIN), Palacky University Olomouc, Slechtitelu 241/27,783 71, Olomouc, Czech Republic. <sup>5</sup>These authors contributed equally: Chad R. Simmons, Tara MacCulloch. <sup>✉</sup>email: [nstepha1@asu.edu](mailto:nstepha1@asu.edu); [hao.yan@asu.edu](mailto:hao.yan@asu.edu)

The fabrication of highly customizable 3D DNA-based architectures for the precise organization of nanoscale materials was originally conceptualized by Seeman in 1982<sup>1</sup>. A variety of methodologies that allow for programmable self-assembly have been developed, including the attachment of DNA linkers to the nanoparticle (NP) surfaces<sup>2,3</sup> for the construction of 3D lattices with user-defined colloidal NP crystal configurations<sup>4–8</sup>. With the advent of DNA origami<sup>9</sup>, 3D superlattices of nanoparticle cluster shapes have been described<sup>10,11</sup> along with 3D origami lattices displaying customizable geometry to host guest species<sup>12</sup>. Rationally designed crystals based upon the “tensegrity” motif<sup>13</sup> which self-assemble with designed four-arm junction crossover points and “sticky end” cohesion have also been shown to be capable of being converted into nanodevices<sup>14</sup>. Recently, several unique motifs with distinctive crystal symmetries, improved resolution, and in one case, a unique junction nucleotide sequence, have been reported<sup>15–18</sup>.

The application of four-way junctions for 3D DNA crystals was inspired by genetic recombination, whereby an unstable branched intermediate termed a Holliday junction (HJ) is created and subsequently undergoes a dynamic reconfiguration to facilitate recombination during cell division<sup>19</sup>. Holliday junctions have been extensively structurally characterized<sup>20–29</sup>, and emerged as a key motif for rationally designed nanoscale assemblies and devices in structural DNA nanotechnology<sup>1,30</sup>. Naturally occurring HJs can “slide” and change the length of their arms<sup>31,32</sup>, a process known as branch migration<sup>33</sup>. Introducing asymmetric sequences at the branching point, however, effectively immobilizes the junction and allows its use in constructing well-defined nanostructures<sup>34</sup>. Although a variety of multi-branch junctions have been employed<sup>35–41</sup>, the four-arm HJ remains the most popular.

Theoretically, there are 36 base pair combinations of immobile sequences. “J1” was the earliest to be designed<sup>42,43</sup>, and has been used nearly exclusively in the construction of self-assembled 3D crystals<sup>13,15–17</sup> with a single exception where “J10” was used<sup>18</sup>. Some early work explored the sequence influence on packing of stacked-X junctions using theoretical<sup>44,45</sup> or experimental methods<sup>46,47</sup>, but to our knowledge, no systematic study of the immobile HJs has been performed. In addition, the relationship between ion concentration and the transition of the HJ from the open to the stacked conformation is well known<sup>20,25,48–51</sup>; however, studies of sequence effects on the ability to bind ions are limited, making the elucidation of the structural parameters that influence crystallization and symmetry a desirable route for exploration.

In this work, we exhaustively probe the ability of all 36 immobile HJ sequences to form DNA crystals in two different designed systems. Our work revealed that a large majority of immobile HJs enable crystallization, with some yielding higher resolution structures and a variety of symmetries. The symmetry is also highly sensitive to the sequence of the junction arms (stems) further away from the branching point, as we demonstrate using scrambled stem sequences. Unique ion binding sites were also observed at two conserved positions within the structures. By performing molecular dynamics (MD) simulations of all 36 HJs in solution, we show that these sites are pivotal for crystallization, with the universally non-crystallizing “fatal” junctions showing zero ability to bind ions in this fashion. Although limited by sampling, the modeling used in this work provides a reliable picture of sequence-dependent flexibility and solvent effects of all 36 immobile HJs as they relate to self-assembled DNA crystal lattices. Overall, the work provides a rigorous and complete description of how different immobile HJ sequences, flanking sequence modifications, and the ability of the junction to capture ions profoundly affect the rational design of self-assembled systems that span a multitude of designer DNA architectures.

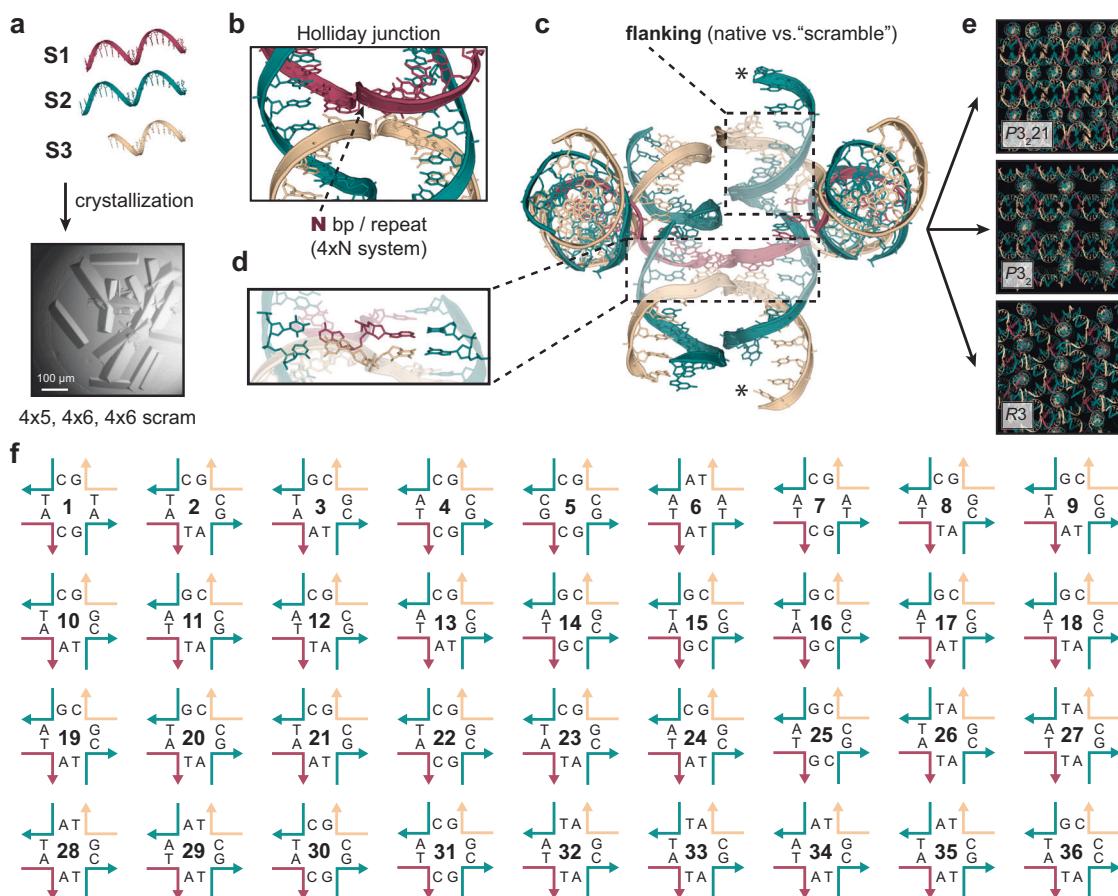
## Results and discussion

Three separate self-assembling DNA crystal systems are described in this report: the “ $4 \times 5$ ”<sup>15</sup> and “ $4 \times 6$ ”<sup>16</sup> designs (collectively referred to as the  $4 \times N$  systems), and a third construct with a “scrambled” sequence variant of the  $4 \times 6$  lattices. Self-assembly was mediated by three constituent oligonucleotides (Fig. 1a): (S1) containing four sequence repeats of either 5 or 6 bases; (S2) composed of 21 bases containing complementary regions to both (S1); and (S3) which forms the second junction crossover (Fig. 1b). Each asymmetric unit could be defined as either a HJ containing 10 and 11 bp on each arm, or as a 21-bp linear duplex (Supplementary Fig. 1). Both versions were solved in parallel, yielding a total of 134 crystal structures (Supplementary Table 1 for data collection and refinement statistics). The crystal lattices contain continuous arrays comprised of a series of crystal “blocks” that self-assemble into a series of 21-bp duplexes tethered by the  $4 \times N$  strand. The HJ serves as the fundamental component at the core of each unit, and the ultimate assembly of the full lattice is facilitated by the complementary 2-base sticky ends that tail each duplex. The “flanking” sequences in the original  $4 \times 6$  were also altered on opposing sides to identify any role the downstream sequence might play in crystallization behavior, independently or in conjunction with the HJ sequence itself.

Here, we probed junction sequence space by creating a panel of 36 immobile HJs (Fig. 1f) in each system while considering a single fixed isomer to determine if junctions other than J1 were capable of crystallizing and potentially yielding improved resolution, and to identify those that proved fatal. The sequences were explicitly defined on each constituent junction strand (all component oligonucleotide sequences can be found in Supplementary Data 1 and Supplementary Table 2), or when defined as a 21-bp duplex (Fig. 1, Supplementary Fig. 2, and Supplementary Fig. 3). To definitively confirm a junction as fatal, rigorous screening (Supplementary Table 3) was performed to conclusively classify each sequence (See “Methods” in the Supplementary Information for technical details). The resulting structures contained three different symmetries ( $P3_221$ ,  $P3_2$ , and  $R3$ ) that are described in detail below (Fig. 1e, Supplementary Fig. 4).

**Junctions enhance resolution and influence crystal symmetry ( $4 \times 5$  motif).** The original  $4 \times 5$  structure containing the J1 junction sequence resulted in a structurally strained array, presumably due to the underwinding to the central component strand<sup>15</sup>. The layered motif provided a well-defined scaffold at 3.1 Å resolution, but the aperiodicity of the cavities and corresponding volumes would be inadequate for scaffolding guest materials. We hypothesized that it might be possible to introduce other unique immobile sequences that could provide slight perturbations in the junction angle, to potentially reduce the strain in the original system and yield a “relaxed,” periodic lattice.

The  $4 \times 5$  system robustly crystallized with 75% of the junctions, and we obtained crystals in all but 9 of 36 sequences (Supplementary Fig. 5 and Supplementary Table 4). Junctions J2 and J30 did crystallize, but were of inadequate quality for structure solution, and therefore classified as fatal. Of the resulting structures, 18 exhibited  $P3_2$  symmetry with average cell dimensions  $a = b = 68.85$  Å  $c = 60.09$  Å, with only nine retaining the original  $P3_221$  symmetry with average cell edges  $a = b = 68.17$  Å  $c = 60.60$  Å (Supplementary Table 5). Although the cell parameters between the two symmetries were virtually indistinguishable (Fig. 2a, Supplementary Fig. 6), the differences between the periodicity of the lattices is dramatic, with vastly different cavity sizes (Fig. 2b, c). While the highest resolution achieved for the J1 system was 3.1 Å, in roughly half of the crystals measured the resolutions were 3.05 Å or better, and as high as 2.9 Å



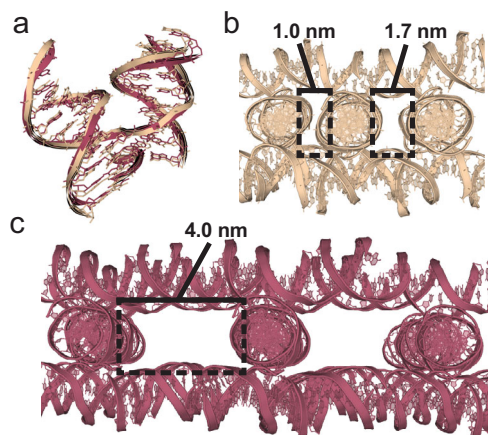
**Fig. 1 Schematic of the composition of the Holliday junction which is required for the self-assembly of 3D DNA lattices.** **a** Three oligonucleotides mediate the crystallization of two self-assembling motifs ( $4 \times 5$  and  $4 \times 6$ ) along with a “scrambled” sequence variant. A representative example of a 3D crystal is shown. **b** The structure of the Holliday junction is the key building block for assembly, and contains four arms using two oligonucleotides (S1; red) and (S3; tan) serving as crossover strands, with S2 (green) serving as a third “linear” complementary strand on each side. The complementary region of S1 contained either five or six bases (N bp) on each arm before each crossover at which point an identical sequence repeats in each consecutive arm for a total of four times ( $4 \times N$ ) before beginning the series again ( $4 \times 5$  or  $4 \times 6$ ). S1 subsequently serves as the scaffolding strand for the entire lattice. **c** The central building “block” that facilitates the 3D assembly. S1 tethers four 21-bp duplexes with the Holliday junction (boxed; translucent) at the core of the structure. The linear 21 base ssDNA oligonucleotide (S2) comprises one half of each duplex with the second crossover strand (S3) flanking each end (boxed). Each duplex is tailed by 2 bp complementary “sticky ends” (asterisks) which cohere to form continuous 3D arrays. **d** Representative bases where sequence asymmetry was imposed to prevent “sliding” of the strands to create 36 immobilized junctions. **e** Three unique symmetries ( $P_{3,21}$ ,  $P_{3,2}$ , and  $R_3$ ) are dictated by the  $4 \times N$  scaffolding strand working in concert with the sequence at each immobile junction. **f** The 36 immobile junction sequences represented in an open Holliday junction format with each strand colored in accordance with (b). Nucleotides on each corresponding strand are indicated with the sequence positions on each component oligonucleotide corresponding to the colored scheme in (d).

(J6) and  $2.75 \text{ \AA}$  (J19), for the  $P_{3,2}$  and  $P_{3,21}$  symmetries, respectively (Supplementary Table 6).

The cavity volumes of the  $P_{3,21}$  lattices were calculated based upon the 3-nm edges of the pores along the three-fold symmetry axis of the crystal with a triangular prism height of 6.1 nm, corresponding to the average  $c$ -axis of the unit cell (Supplementary Fig. 7). The aperiodic lattice contained cavities with widths of 1.0 and 1.7 nm at alternating intervals, and yielded exceedingly small pore volumes of  $\sim 24 \text{ nm}^3$  (Fig. 2b; Supplementary Fig. 7a). The  $P_{3,2}$  channels were treated as a hexagonal prism with 6.4 nm along each edge and a height measured from the top to the bottom of each block along with the average  $c$ -axis = 6.0 nm (Supplementary Fig. 7b), with a resulting volume of  $\sim 639 \text{ nm}^3$ , a nearly 27-fold increase in cavity size relative to the  $P_{3,21}$  lattices. The junction PDB coordinates for each symmetry type were grouped, and their angles were analyzed in DSSR<sup>52</sup>. The mean angles across all junctions within the groups were  $56.05^\circ$  ( $\sigma = 1.63$ ) and  $56.59^\circ$  ( $\sigma = 1.50$ ) for the  $P_{3,2}$  and  $P_{3,21}$ , respectively (Supplementary Table 7). We hypothesize that even

a small difference in junction angles could have a significant global effect on the assembly of the lattice, thereby eliminating the strain in the  $P_{3,21}$  lattices.

**Junction sequences can dramatically alter crystal symmetry ( $4 \times 6$  motif).** In parallel, we employed the previously reported J1  $4 \times 6$  system ( $3.05 \text{ \AA}$ ), with  $P_{3,2}$  symmetry<sup>16</sup>. All 36 junction sequences were screened, and consistently fatal junctions were identified. Seventeen of the 36 junctions successfully crystallized, a 47% success rate (Supplementary Fig. 8 and Supplementary Table 8). Unlike the  $4 \times 5$  motif, nearly all of the  $4 \times 6$  junction sequences had a strong preference for crystallization in buffers containing  $\geq 2.0 \text{ M}$  salts (e.g., LiCl,  $\text{Li}_2\text{SO}_4$ , KCl, and NaCl), with the majority preferring slightly basic pH conditions in cacodylic acid (Supplementary Table 9). We saw no appreciable improvements in resolution; however, in five cases (J4, 5, 31, 33, and 36), the junction altered the symmetry from trigonal ( $P_{3,2}$ ) to rhombohedral ( $R_3$ ). Furthermore, J4 and J36



**Fig. 2 Discrete differences in junction angles determine global lattice symmetry.** **a** Superimposed structures of the J5 (tan) and J3 (red)  $4 \times 5$  junctions containing  $P_{3_2}21$  and  $P_{3_2}$  symmetries, respectively. The junction alignment had a global RMSD value of 1.34, with calculated interduplex angles of  $58.18^\circ$  and  $55.20^\circ$ , respectively. No significantly obvious visual differences are apparent; however, the resulting global influence that an even modest difference in angle can have on overall packing is evident in **(b, c)**. **b** Snapshot of the full J5  $P_{3_2}21$  ( $4 \times 5$ ) lattice containing an aperiodic array of cavities which would not be amenable for scaffolding of guest molecules of any appreciable size. The two uniquely sized cavities are shown with black boxes. The widths for each respective cavity are indicated. Each cavity spans the length of a cross-section of a duplex ( $\sim 2.0$  nm). Alternative views of the lattice including measurements in each orientation are included in Supplementary Fig. 8. **c** Snapshot of the full J3  $P_{3_2}$  ( $4 \times 5$ ) lattice which reveals dramatically different arrays of large periodic cavities compared to its J5 counterpart in **(b)**. A single cavity is highlighted with a black box with a width of 4.0 nm and also spanning the length of a cross-section of a duplex ( $\sim 2.0$  nm). Alternative views of the lattice including measurements in each orientation are included in Supplementary Fig. 8.

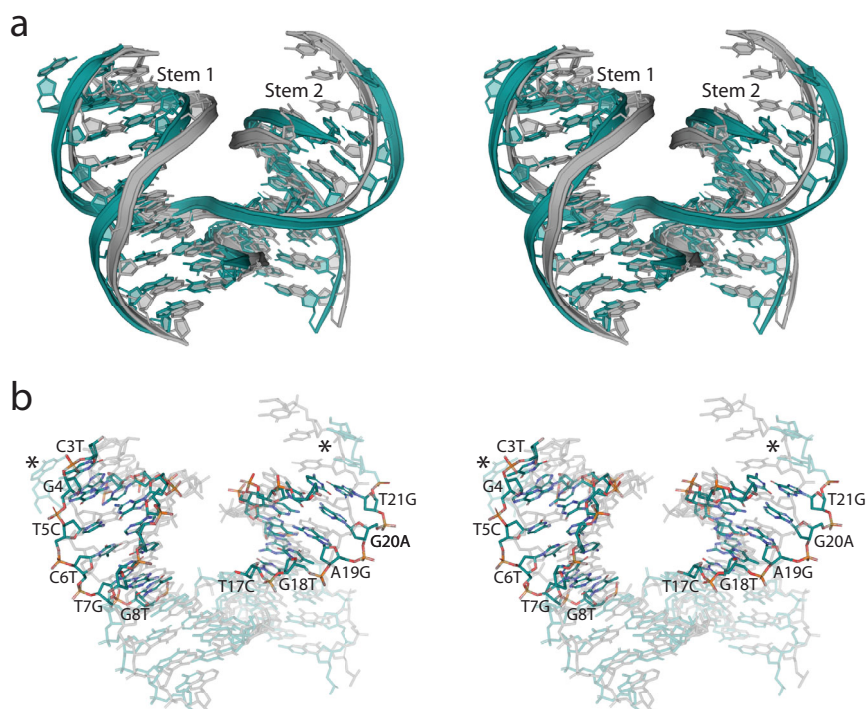
crystallized exclusively in  $R3$ , but J5, 31, 33 exhibited the ability to crystallize in both  $P_{3_2}$  and  $R3$  (Supplementary Fig. 9). In these scenarios,  $R3$  preferred low concentrations of divalent ions and organic solvents, and  $P_{3_2}$  required high salt (Supplementary Table 10) with drastically different lattices (Supplementary Fig. 10). When compared to the junctions that did not yield crystals in the  $4 \times 5$  system, six (J11, 12, 13, 17, 18, and 27) junctions consistently proved fatal (Supplementary Table 11), and we suggest that these sequences could be considered unwise options for future design decisions.

The average cell constants for the  $P_{3_2}$  crystals were  $a = b = 68.29 \text{ \AA}$   $c = 55.68 \text{ \AA}$ . The  $c$ -axis, in particular, was  $\sim 5 \text{ \AA}$  shorter, and had a much larger degree of variability ( $\sigma = 1.97$ ), than those that crystallized using the  $4 \times 5$  motif (Supplementary Table 12). The broad range of the short axis spanned from 52.77 to 60.36  $\text{\AA}$ , and we posit that the flexible axis lengths could be responsible for rendering a larger number of junctions fatal than the  $4 \times 5$  system. In crystals containing  $R3$  symmetry, the average cell for each was  $a = b = 114.9 \text{ \AA}$   $c = 49.77 \text{ \AA}$  with the  $c$ -axis confined to a tighter regime ( $\sigma = 0.75$ ). The cavity volumes were also markedly different from the  $P_{3_2}$  cavity volumes ( $\sim 614.7 \text{ nm}^3$ ) compared with  $\sim 532.1 \text{ nm}^3$  in the more densely packed  $R3$  structure (see Supplementary Fig. 11 for calculation details). The average junction angles of  $54.60^\circ$  ( $\sigma = 1.44$ ) and  $58.37^\circ$  ( $\sigma = 2.3$ ) for the  $P_{3_2}$  and  $R3$  symmetries, respectively (Supplementary Table 13), along with their preference for salts or organic solvents, were likely the major contributing factors that yielded the divergent lattices.

**Role of the junction flanking sequence on symmetry in the  $4 \times 6$  motif.** Because the  $4 \times 6$  system yielded  $R3$  symmetry in 5 of 17 junctions, we considered whether the downstream sequences adjacent to the junction could play a role in crystallization efficiency, junction angle, and symmetry preference, or if the HJ alone was the singular determinant. To investigate this possibility, we designed “scrambled” sequences with targeted base substitutions (Supplementary Fig. 12), while maintaining GC content, along each “stem” (Fig. 3a, b). Contrary to the buffer preference observed with the native  $P_{3_2}$  crystals, the scrambled systems exhibited an exclusive preference towards low salt buffers (Supplementary Fig. 13 and Supplementary Table 14), much like the native  $R3$  crystals. Remarkably, only J1 and J2 retained the  $P_{3_2}$  symmetry, with all others yielding an  $R3$  lattice (Supplementary Table 15). It is also noteworthy that J1 and J2 ( $P_{3_2}$ ), along with the  $R3$  systems, also shared a preference for low salt conditions, unlike the original  $P_{3_2}$  crystals. This change in buffer preference suggests that global sequence content can indeed influence self-assembly behavior. Further, we observed modest improvements in resolution, reaching as high as 2.7  $\text{\AA}$  in J36. Notably, all junctions preventing crystallization compared to its original counterpart remained fatal (Supplementary Table 16) and all relative cavity dimensions and volumes remained unperturbed (Supplementary Figs. 14 and 15).

The role of longer-range sequence effects on junction angles remains a largely open question, but the effect in our crystal system is significant. There appeared to be only a marginal difference in the average cell lengths in the  $R3$  scramble crystals ( $a = b = 113.04$   $c = 51.10$ ; Supplementary Table 17) relative to the native  $4 \times 6$  sequences (Supplementary Table 12): the  $a$ ,  $b$  axis trended  $\sim 2 \text{ \AA}$  shorter and  $c \sim 1.3 \text{ \AA}$  longer, whereas in the J1 and J2 cases, the crystals were in good agreement with the original  $4 \times 6$  motif. The average angles of the  $R3$  and  $P_{3_2}$  crystals were  $61.00^\circ$  ( $\sigma = 1.21$ ) and  $58.05^\circ$  ( $\sigma = 1.39$ ), respectively. Although the average angle in the  $R3$  crystals is nearly  $3^\circ$  higher than those from the native sequence (Supplementary Table 18), the scrambled sequence angle is calculated from a larger sample size than the smaller angle calculated in the native  $4 \times 6$  structures ( $58.37^\circ$ ), with a significantly larger standard deviation (2.33). By contrast, the calculated angle for the  $P_{3_2}$  crystals was  $58.05^\circ$  ( $n = 2$ ,  $\sigma = 1.39$ ), compared with  $54.60^\circ$  ( $n = 16$ ,  $\sigma = 1.44$ ), a less accurate average due to a small sample size, so we posit that  $54.60^\circ$  more closely reflects the average observed angles in the  $4 \times 6$  systems.

**Conserved ion binding sites within the junction structure influence crystallization.** We previously reported<sup>15</sup> the presence of arsenic ions at two opposing positions (Pos1 & 2, Supplementary Figs. 16 and 17) of the junction, due to the cacodylic acid contained in the crystallization buffer (Fig. 4). In a number of cases, we also observed a clustering of ions (Pos3; Supplementary Fig. 17) within the minor groove neighboring Pos2, but shared no apparent interactions with the junction. Pos1 and 2 were readily observable in the electron density maps at either one or both of these conserved sites in a significant number of the structures, with no apparent respect to motif or symmetry (Fig. 4). Although in the majority of cases, the individual junctions, regardless of design parameters, had a preference for crystallizing in cacodylic acid within the 6.0–6.5 pH regime, not all crystals were confined to this requirement. In the  $4 \times 5$  system, J25 and J34, both displaying  $P_{3_2}$  symmetry, crystallized in 50 mM Tris pH = 8.0 buffer containing cobalt hexamine ( $\text{CoH}_{18}\text{N}_6$ ) and 10 mM  $\text{MgCl}_2$ , and 50 mM HEPES pH = 7.5 with 20 mM  $\text{MgCl}_2$ , respectively. Only cobalt could account for the different peaks in the  $F_o - F_c$  maps in J25, whereas in J34, and J22 and J23 in the  $4 \times 6$  and



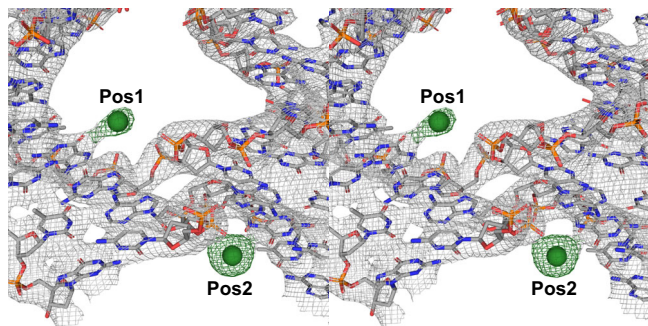
**Fig. 3 Stem sequence perturbations alter junction angles and influence global symmetry.** **a** Stereoview of a superposition of the J10 junction structure using the original  $4 \times 6$  sequence motif with  $P3_2$  symmetry (gray), with the scrambled sequence version containing  $R3$  symmetry (teal). The modified sequences were located within the two downstream stem (1 & 2; indicated) regions containing the same GC content as the original  $4 \times 6$  sequence version. The effect of the scrambled sequence on the geometry of the junction is visually apparent when comparing the superimposed Stem 1 & 2 regions. The dramatic difference in junction angle, and its influence on symmetry is evident (Supplementary Fig. 15). **b** Stereoview of a stick representation of the superimposed J10 structures in **(a)** with all base modification sites between the original and scrambled sequence version of the  $4 \times 6$  J10 structures are indicated in Stems 1 & 2. Asterisks are included to provide attention towards the sticky end regions that significantly diverge as an apparent result of the angles induced by the modified stem sequences. Atoms are indicated using the following: carbon (teal), nitrogen (blue), oxygen (red), and phosphate (orange). All regions containing identical sequences are left translucent.

$4 \times 6$  scramble systems, respectively, contained  $Mg^{2+}$ . The modeled ions were readily superimposable, and well-coordinated at Pos1 & 2 (Supplementary Fig. 18).

**Molecular dynamics simulations of individual HJs reveal interhelical angle differences.** Full-atomistic MD simulations were performed for all 36 immobile HJs to explore their dynamics in solution and compare the properties of the crystallizing and non-crystallizing junctions (see the “Methods” section in the Supplementary Information for full technical details). In over 224  $\mu s$  of simulations, the force-field performance was deemed satisfactory, as stable base pairing and B-form helical topologies were observed in agreement with our crystallographic experiments and earlier studies<sup>28,53–55</sup>. The simulations revealed relatively minor differences in the interhelical dynamics among the 36 immobile HJs. We present the median angle values, as well as histograms of angle populations in Supplementary Table 19 and Supplementary Fig. 18, respectively. The interhelical angle rapidly fluctuated on the microsecond timescale, reflecting the greater conformational freedom of the HJs in solution. Its median value was typically lower than the one seen in the self-assembled DNA crystals reported in this work, likely reflecting a genuine influence of the respective environments (crystal lattice vs. free solution), as the simulation values are more consistent with those reported for X-ray structures of isolated HJs<sup>56</sup>. Still, for the majority of the HJs the difference was less than  $5^\circ$ . We note that the most unusual

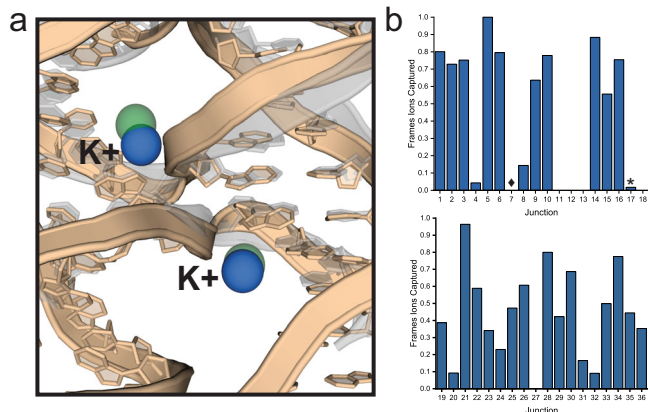
interhelical angle values and distributions were observed for the J11 and J18 junctions, both of which never crystallized. The excessive interhelical dynamics and angle preferences incompatible with the lattice structure could be a factor contributing to the crystal growth inhibition for these two specific junctions.

**The ion binding ability of the junction could determine its potential to form 3D lattices.** The most significant difference between the individual junctions observed in MD simulations was their ability to form distinct potassium ion binding sites near the junction branching point. In all cases, the ion formed a bridge between the phosphate right at the branching point and one or two closest bases. These sites were in good agreement with the Pos1 and Pos2 sites observed in the experimental crystal structures (Figs. 4 and 5a). Since the base atoms of the branching point base pairs were involved in the ion coordination, the ion-binding sites were highly different among the 36 immobile HJs. The most obvious and striking difference was that the HJs which never crystallized in our experiments (J11, 12, 13, 18, and 27) were also those which consistently showed no ability to form these specific ion-binding sites in simulations. J17, which also never crystallized, did so in a negligible portion (0.02%) of all simulation frames (Fig. 5b). All the other HJs both crystallized in our experiments and formed these ion binding sites in our simulations to some degree (Fig. 5b). The mean incidence of binding in the non-fatal junctions was 0.53 ( $\sigma = 0.28$ ), suggesting that the



**Fig. 4 The junction crossovers contain unique ion binding positions.**

Stereoscopic view, using J21 in the  $4 \times 5$  system,  $2F_o - F_c$  electron density accounting for the bases at the crossover regions are contoured at  $\sigma = 2.0$ , and the individual ion positions 1 and 2 (indicated Pos1 and Pos2) are independently contoured in the corresponding electron density at  $\sigma = 4.0$ . The presence of arsenic at these sites was substantiated by transferring the crystals into TAE- $Mg^{2+}$  (40 mM Tris, 20 mM acetate, and 1 mM EDTA pH = 8.6), and subsequently freezing the crystals. The crystals were scanned at the arsenic K edge ( $\lambda = 1.04 \text{ \AA}$ ) where the corresponding arsenate peak was present. No other components within the crystallization buffers could account for the resulting peaks in the  $F_o - F_c$  difference maps for the ions at their corresponding sites. Atoms are indicated using the following: carbon (gray), nitrogen (blue), oxygen (red), phosphate (orange), and arsenic ions (green spheres).



**Fig. 5 MD simulations relate capture of ions with lattice formation.**

**a** Superimposed structures of the  $4 \times 5$  J10 crystal structure (translucent gray) with a snapshot from the J10 simulation in solution (tan). The arsenic binding sites (green spheres) in the crystal structure near the branching point overlap with spontaneously formed potassium (blue spheres) binding sites in the solution structure simulations. **b** Graphs showing the incidence of ion capture near the branching point in simulations of all 36 junction sequences. The consensus “fatal” junctions (J11, 12, 13, 17, 18, and 27) show no ability for ion capture with the exception of J17 (asterisk) which did so to a negligible degree compared to the crystallizing junctions. All other junctions resulting in crystals demonstrated the ability for ion capture to a significant degree, with only a single outlier (J7, diamond). J7 robustly crystallized, but showed no ability to capture ions in both experiments and simulations, suggesting the ion binding is not essential for crystallization of this single junction.

ability to capture ions is pivotal to the ability to crystallize the immobile HJs. We speculate that the branching-point ion-binding site could be stabilizing the DNA strand exchange during the formation of the lattice and thus facilitate the crystal growth. This hypothesis is supported by the fact that the junctions unable to form this ion-binding site, or those which do so to a lesser degree,

are consequently among those that either do not grow crystals at all, or that are more sensitive to the crystallization conditions. The sole exception was J7 which, despite possessing an appropriate branching point sequence to do so, did not form the ion-binding site in simulations or experiments, but was still able to crystallize.

Initially, there might be an apparent contradiction associated with the experimentally detected presence of arsenic from a cacodylate anion in the region where the simulations localize  $K^+$  cations (Fig. 5). However, we suggest that these observations can be reconciled. To our knowledge, all experimental structures of HJs in the PDB database reported by diverse independent groups suggest that cations (e.g.,  $Na^+$ ,  $Mg^{2+}$ ,  $Sr^{2+}$ ,  $Ba^{2+}$ ) should bind at this location. This result is consistent with the highly negative molecular interaction potential (Supplementary Fig. 20) calculated here, which is a hallmark of cation-binding sites. However, in the majority of the crystal structures reported here, the cacodylate anion indeed fits best into the electron density map of our X-ray structures, due to the presence of sodium cacodylate in the crystallization solution (Supplementary Fig. 21). No other buffer component could form any reasonable contact distances at this location with the junction itself. A complete description of the comprehensive, experimental evidence for this arsenic has been described in our previous work<sup>15</sup>. This seemingly counter-intuitive result can be rationalized in several ways: First, the cacodylate could be stabilized at this site by hydrogen bonding and interactions with the solvent, both of which are known to provide stabilization for anionic binding even at regions with negative surface potentials<sup>57,58</sup>. Second, there could be one or more sodium counterions associated (catenated) with cacodylate at this site. Cacodylate anions are known to form interactions with negatively charged segments of nucleic acids in this fashion<sup>59,60</sup>. Lastly, there could be multiple  $Na^+$  ions (the counterion of cacodylate) present around the anion. Due to the nominal resolution of our structures, which is in the  $\sim 3 \text{ \AA}$  range, it is not possible to unambiguously identify all the interacting species, nor is it possible to fully describe the exact coordination, such as the involvement of water molecules and  $Na^+$ . The binding of  $Na^+$  could compensate or even over-compensate the cacodylate negative charge, effectively re-creating the common cation-binding site seen in simulations and other experimental structures. The  $Na^+$  ions are often fluctuating and their coordination requirements are pliable, rendering them entirely invisible to us. Furthermore, as suggested by simulations, there could be significant local dynamics that obscure the densities with averaging. Past MD simulations reported highly variable  $Na^+$  binding sites, from stiff to dynamic ones, around nucleic acids<sup>61</sup>, and it is entirely possible that invisible  $Na^+$  cations can bridge the cacodylate anion with the DNA carbonyl and phosphate groups. While the inability to describe the interaction site in greater detail is a limitation of the current work, it is abundantly clear that accommodation of ions at this location is a strict requirement for the junction to stabilize, and to effectively crystallize.

In this work, we have reported a systematic study of all 36 immobile HJ sequences across two different crystal systems, findings that can, in principle, be extended to any 3D DNA crystal and potentially other types of DNA architectures. We show that J1 (or any other junction) should not be considered a privileged option for designing self-assembled lattices. Rather, a host of other sequence combinations discussed herein should be explored, with many of them providing potentially superior performance. We make the essential observation that several junctions—including J11, 12, 13, 17, 18, and 27—are universally fatal to crystallization and should be avoided in future crystal designs, unless perhaps alternate isomers are considered. Another major observation is the importance of stem region sequences

outside the HJ, given that this sequence can control crystal symmetry and lattice architecture, and improve resolution in non-obvious ways. We also elucidated the key role of ion coordination in driving these effects. Scaling up DNA nanostructures to larger assemblies—for DNA crystals, as well as 1D and 2D lattices, and potentially DNA origami systems—will likely require taking into account sequence-dependent effects both at the level of stem and HJ geometries. Finally, the angle distributions obtained from our MD studies of junctions can improve accuracy of coarse-grained models and DNA nanotechnology design tools to more accurately represent nanostructures.

In summary, our experimental results are corroborated by the largest scale MD (224- $\mu$ s-aggregate) simulation performed to date on HJs. The simulations elucidated the role of sequence-dependent flexibility and solvent effects on HJ conformations. This work also demonstrates that HJ sequence plays a non-trivial role in the ability of DNA crystals to form, and can dramatically influence crystal symmetry. Furthermore, we have provided a systematic and comprehensive sequence-structure study on self-assembled DNA crystal systems using all 36 junctions. Lastly, the study unexpectedly revealed specific molecular interactions (ion binding) using both experimental and modeling parameters, that has never been performed on this scale.

## Methods

**Crystallization and data collection.** All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa) and purified using 14% denaturing polyacrylamide gel electrophoresis (PAGE) or HPLC. Following purification, the pelleted DNA was resuspended in nanopure H<sub>2</sub>O and washed 5 $\times$  using 3 kDa molecular weight cut off filters (Amicon) to remove any remaining salt. A stock of the three-component strands (S1:S2:S3) was made with final concentrations of 30:120:120  $\mu$ M for each construct. The sitting drop vapor diffusion method was performed with Cryschem plates (Hampton Research) using an adaptation of a discontinued DNA crystallization screen from a commercial vendor (Sigma Aldrich) containing a sparse matrix of 48 conditions (Supplementary Table 3). 500  $\mu$ L of each corresponding condition was added to each reservoir, and a total drop volume of 6  $\mu$ L was prepared containing a mixture of a 2:1 ratio of DNA stock to the corresponding reservoir solution with a final drop concentration of 30  $\mu$ M. For junctions where crystallization was challenging, multiple rounds of screening using the 48-condition sparse matrix screen (Supplementary Table 3) were performed by varying conditions such as DNA concentration, buffer pH, salt concentration, and annealing times to allow for a reliable determination of the efficacy of the junction. The plates were then placed in a chilling incubator (Torrey Pines Scientific, Carlsbad, CA), and allowed to equilibrate to 60 °C for 1 h and then cooled using a linear gradient to 25 °C at a rate of 0.3 °C/h. The resulting crystals were imaged using a light microscope (Supplementary Figs. 6, 9, 10), and then cryo-protected using an artificial mother supplemented with 30% glycerol by direct addition to the drop. The crystals were subsequently harvested using cryo-loops (Hampton Research) and cryo-cooled by immediate submersion into a liquid nitrogen bath. All data were collected in a nitrogen cold-stream (100 K) at the corresponding beamlines indicated in Supplementary Table 1.

**Data processing and structure solution.** All diffraction data were processed in HKL2000<sup>62</sup>, and the initial phases were calculated using molecular replacement in Phaser<sup>63</sup> from the PHENIX<sup>64</sup> suite of programs with either the J1 4  $\times$  5 structures 5KEK and 6X8C, as the initial search models for the duplex and junction structures, respectively, and the J1 4  $\times$  6 structures 5VY6 and 6XNA for the duplex and junction structures containing P<sub>3</sub> symmetry in that system, respectively. However, with few exceptions, the majority of the 4  $\times$  6 R3 symmetry crystals required another R3 model of one of its counterparts as the ideal search model. Multiple rounds of model building were performed in Coot, with the initial model first treated as a single rigid body, followed by subsequent iterative rounds using restrained refinement in REFMAC<sup>65</sup> from CCP4<sup>66</sup>, along with real space, and XYZ coordinate calculation in phenix.refine. All ions were modeled into regions of F<sub>o</sub>-F<sub>c</sub> difference density with a contour level  $\geq \sigma = 3.0$  and refined. Atom occupancies and B-factor calculations were then refined, along with simulated annealing to conclude refinement. All refined models used an R<sub>free</sub> set containing 5–10% of the unique reflections for each structure. The coordinates and structure factors, totaling 134 unique structures, were deposited in the Protein Data Bank (PDB), and the corresponding accession codes are listed in Supplementary Table 20. Data collection and refinement statistics are all summarized in Supplementary Table 1 which is divided according to their respective crystal systems. All main text and Supplementary Information structure figures contained in this report were prepared using PyMOL<sup>67</sup>.

**Molecular dynamics simulations.** We have used the structure of the junction J1 crystallized in 4  $\times$  5 lattice geometry (PDB:5KEK)<sup>15</sup> as the starting structure for all simulations. Each arm of the HJ was extended to contain at least eight base pairs, with the extended parts of the helices taken from the adjacent cells of the crystal lattice. The resulting construct contained 64 nucleotides and was used as the initial structure for simulations of all the 36 immobile junction sequences, by substituting the branching point base pairs accordingly. We have used xLeap module of AMBER18<sup>68</sup> to prepare the topology and coordinate files. The latest AMBER OL15 DNA force field<sup>69</sup> was used in all reported simulations. During the revision, we also tested the alternative parmbsc1 DNA force field<sup>70</sup> for a subset of the HJs; however, with bsc1 we detected a significant population of non-native and possibly spurious  $\beta/\gamma$  g + /t backbone conformations (see Supplementary Fig. 22 and the accompanying text (Supplementary Discussion 1)). This observation is also consistent with recent reports<sup>71</sup>. We thus suggest that for the present system the OL15 force field might be the optimal choice. Each HJ structure was solvated in an octahedral box of SPC/E water molecules<sup>72</sup> with a minimal distance of 16 Å between the solute and the box border (Supplementary Fig. 23). The 0.15 M salt concentration was established by the addition of KCl ions<sup>73</sup>. The relative ion positions were then compared to the ion sites in the crystal structures. Although the MD conditions and description of ion binding are not identical to the experimental conditions, the simulations should quite realistically reflect the relative overall propensities of different sequences to form the ion-binding site. Next, we performed pre-production equilibration and minimization of each system. The first minimization was performed with positional restraint of 25 kcal/mol/Å<sup>2</sup> placed on the DNA followed by an equilibration run using the same positional restraint in which the system was heated from 100 to 300 K on the timescale of 10 ps. This was then followed by a series of six minimizations and equilibration runs with 5, 4, 3, 2, 1, and finally, 0.5 kcal/mol/Å<sup>2</sup> positional restraint placed on the DNA. Each minimization consisted of 500 steps using the steepest descent method followed by 500 steps using the conjugate gradient method. Each equilibration, except the first one, was performed for 5 ps. For all junctions other than J1, an extra minimization step was taken at the beginning to optimize the initial geometry of the branching point base pairs. Production simulations were performed with the pmemd.cuda<sup>74</sup>, using periodic boundary conditions and NPT ensemble, and the standard simulation protocol<sup>75</sup>. The length of each simulation was 1  $\mu$ s and four independent simulations were performed for all 36 junctions. Simulations of selected HJs were then extended up to 20  $\mu$ s to verify the convergence (Supplementary Fig. 24a, b). The analyses were performed with cpptraj and VMD<sup>76,77</sup>, using the combined simulation ensemble of each individual junction. The interhelical angles of the junctions in simulations were measured as J<sub>twist</sub> parameters according to the definition previously described by Watson et al.<sup>56</sup> where the two helical axes of the stacked helical arms of the junction are represented by a vector and the J<sub>twist</sub> calculated as their dot angle product. Note that our simulations sampled transitions between right-handed and left-handed junctions which, however, have the same J<sub>twist</sub> values assigned by the Watson et al. 2004 definition. To differentiate the handedness, we defined an additional vector perpendicular to the junction branching point and then used it to calculate the dot angle product with the cross-product vector of the two vectors representing the helical axes. The value of this second dot angle was then used to differentiate the handedness of the junction with values above and below 90° corresponding to right-handed and left-handed junctions, respectively. The J<sub>twist</sub> values of the left-handed structures were subsequently normalized by  $-1$ .

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All coordinates and structure factors generated in this study have been deposited in the RCSB Protein Data Base under the following accession codes (4  $\times$  5 duplex structures: J1-5KEK, J3-6WQG, J5-6WRB, J6-6X8B, J7-6WSN, J8-6WSO, J9-6WSP, J10-6WSQ, J14-6WSR, J15-6WSS, J16-6WST, J19-6WSU, J20-6WSV, J21-6WSW, J22-6WSX, J23-6WSY, J24-6WSZ, J25-6WT0, J26-6WRJ, J28-6WRI, J29-6WT1, J31-6WRC; 4  $\times$  5 junction structures: J1-6X8C, J3-6XDV, J5-6XDW, J6-6XDX, J7-6XDY, J8-6XDZ, J9-6XEI, J10-6XEJ, J14-6XEK, J15-6XEL, J16-6XEM, J19-6XFC, J20-6XFD, J21-6XFE, J22-6XFF, J23-6XFG, J24-6XFW, J25-6XGM, J26-6XFX, J28-6XFY, J29-6XFF, J31-6XGO, J32-6XGJ, J33-6XGN, J34-6XGO, J35-6XGK, J36-6XGL; 4  $\times$  6 duplex structures with P<sub>3</sub> symmetry: J1-5VY6, J2-7JPB, J5-7JPA, J7-7JPC, J8-7JPD, J9-7JPE, J10-7JPF, J16-7JPG, J22-7JPS, J23-7JON, J24-7JOL, J26-7JOK, J28-7JOJ, J30-7JOI, J31-7JOH, J33-7JOG; 4  $\times$  6 junction structures with P<sub>3</sub> symmetry: J1-6XNA, J2-7JFT, J5-7JFU, J7-7JFV, J8-6XO5, J10-7JFW, J16-7JFX, J20-7JH8, J22-7JH9, J23-7JHA, J24-7JHB, J26-7JHC, J28-6XO6, J30-6XO7, J31-6XO8, J33-6XO9; 4  $\times$  6 duplex structures with R3 symmetry: J4-7JRY, J5-7JRS, J31-7JSD, J33-7JSI, J36-7JS2; 4  $\times$  6 junction structures with R3 symmetry: J4-7JHR, J5-7JHS, J31-7JHT, J33-7JHU, J36-7JHV; 4  $\times$  6 scramble duplex structures: J1-7JKD, J2-7JKE, J3-7JKG, J5-7JKH, J7-7JKI, J8-7JKJ, J10-7JKK, J14-7JL9, J16-7JLA, J19-7JLB, J21-7JLC, J22-7JLD, J23-7JLE, J24-7JLF, J26-7JNJ, J30-7JSB, J31-7JSC, J33-7JNK, J34-7JNL, J36-7JNM; 4  $\times$  6 scramble junction structures: J1-7JKO, J2-7JLZ, J3-7JYJ, J5-7JYX, J7-7JYV, J8-7JJE, J10-7JJS, J14-7JJ4, J16-7JJ3, J19-7JJ2, J21-7JIQ, J22-7JIP, J23-7JIO, J24-7JIN, J26-7JIM, J30-7J19, J31-7J18, J33-7J17, J34-7J16, J36-7J15). In addition, all corresponding

accession codes can be found in Supplementary Table 20. The input and output files for MD simulation data generated in this study have been deposited in the Zenodo database under accession code: [6381939](https://doi.org/10.5281/zenodo.6381939). The raw MD simulation trajectory data are available upon request due to the large size of the datasets.

Received: 27 October 2021; Accepted: 4 May 2022;

Published online: 03 June 2022

## References

- Seeman, N. C. Nucleic-acid junctions and lattices. *J. Theor. Biol.* **99**, 237–247 (1982).
- Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382**, 607–609 (1996).
- Alivisatos, A. P. et al. Organization of ‘nanocrystal molecules’ using DNA. *Nature* **382**, 609–611 (1996).
- Seeman, N. C. & Gang, O. Three-dimensional molecular and nanoparticle crystallization by DNA nanotechnology. *Mrs Bull.* **42**, 904–912 (2017).
- Laramy, C. R., O’Brien, M. N. & Mirkin, C. A. Crystal engineering with DNA. *Nat. Rev. Mater.* **4**, 201–224 (2019).
- Park, S. Y. et al. DNA-programmable nanoparticle crystallization. *Nature* **451**, 553–556 (2008).
- Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. DNA-guided crystallization of colloidal nanoparticles. *Nature* **451**, 549–552 (2008).
- Laramy, C. R. et al. Controlled symmetry breaking in colloidal crystal engineering with DNA. *ACS Nano* **13**, 1412–1420 (2019).
- Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- Majewski Pawel, W. et al. Resilient three-dimensional ordered architectures assembled from nanoparticles by DNA. *Sci. Adv.* **7**, eabf0617 (2021).
- Tian, Y. et al. Ordered three-dimensional nanomaterials using DNA-prescribed and valence-controlled material voxels. *Nat. Mater.* **19**, 789–796 (2020).
- Zhang, T. et al. 3D DNA Origami Crystals. *Adv. Mater.* **30**, <https://doi.org/10.1002/adma.201800273> (2018).
- Zheng, J. P. et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461**, 74–77 (2009).
- Hao, Y. D. et al. A device that operates within a self-assembled 3D DNA crystal. *Nat. Chem.* **9**, 824–827 (2017).
- Simmons, C. R. et al. Construction and structure determination of a three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *J. Am. Chem. Soc.* **138**, 12690–12690 (2016).
- Simmons, C. R. et al. Tuning the cavity size and chirality of self-assembling 3D DNA crystals. *J. Am. Chem. Soc.* **139**, 11254–11260 (2017).
- Zhang, F., Simmons, C. R., Gates, J., Liu, Y. & Yan, H. Self-assembly of a 3D DNA crystal structure with rationally designed six-fold symmetry. *Angew. Chem.-Int. Ed.* **57**, 12504–12507 (2018).
- Simmons, C. R. et al. A self-assembled rhombohedral DNA crystal scaffold with tunable cavity sizes and high-resolution structural detail. *Angew. Chem. Int. Ed.* **59**, 18619–18626 (2020).
- Holliday, R. A mechanism for gene conversion in fungi (reprinted). *Genet. Res.* **89**, 285–307 (2007).
- Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Ho, P. S. The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl Acad. Sci. USA* **97**, 3971–3976 (2000).
- Ho, P. S. Structure of the Holliday junction: applications beyond recombination. *Biochem. Soc. Trans.* **45**, 1149–1158 (2017).
- Shida, T., Iwasaki, H., Shinagawa, H. & Kyogoku, Y. Characterization and comparison of synthetic immobile and mobile Holliday junctions. *J. Biochem.* **119**, 653–658 (1996).
- Yadav, R. K. & Yadava, U. Molecular dynamics simulation of hydrated d(CGGGTACCCG)(4) as a four-way DNA Holliday junction and comparison with the crystallographic structure. *Mol. Simul.* **42**, 25–30 (2016).
- Starr, F. W. et al. Holliday junction thermodynamics and structure: comparisons of coarse-grained simulations and experiments. *Biophys. J.* **110**, 178A–178A (2016).
- McKinney, S. A., Declais, A. C., Lilley, D. M. J. & Ha, T. Structural dynamics of individual Holliday junctions. *Nat. Struct. Biol.* **10**, 93–97 (2003).
- Weiss, S. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat. Struct. Biol.* **7**, 724–729 (2000).
- Clegg, R. M. et al. Fluorescence resonance energy-transfer analysis of the structure of the 4-way DNA junction. *Biochemistry* **31**, 4846–4856 (1992).
- Yu, J., Ha, T. J. & Schulten, K. Conformational model of the Holliday junction transition deduced from molecular dynamics simulations. *Nucleic Acids Res.* **32**, 6683–6695 (2004).
- Eis, P. S. & Millar, D. P. Conformational distributions of a 4-way DNA junction revealed by time-resolved fluorescence resonance energy-transfer. *Biochemistry* **32**, 13852–13860 (1993).
- Pinheiro, A. V., Han, D. R., Shih, W. M. & Yan, H. Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.* **6**, 763–772 (2011).
- Hsieh, P. & Panyutin, I. G. In *Nucleic Acids and Molecular Biology* (eds Eckstein, F. & Lilley, D. M. J.) 42–65 (Springer Berlin Heidelberg, 1995).
- Seeman, N. C. & Kallenbach, N. R. DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53–86 (1994).
- Holliday, R. Molecular aspects of genetic exchange and gene conversion. *Genetics* **78**, 273–287 (1974).
- Zhang, S. W., Fu, T. J. & Seeman, N. C. Symmetrical immobile dna branched junctions. *Biochemistry* **32**, 8062–8067 (1993).
- Ma, R. I., Kallenbach, N. R., Sheardy, R. D., Petrillo, M. L. & Seeman, N. C. Three-arm nucleic acid junctions are flexible. *Nucleic Acids Res.* **14**, 9745–9753 (1986).
- Wang, X. & Seeman, N. C. Assembly and characterization of 8-arm and 12-arm DNA branched junctions. *J. Am. Chem. Soc.* **129**, 8169–8176 (2007).
- Wang, Y. L., Mueller, J. E., Kemper, B. & Seeman, N. C. Assembly and characterization of five-arm and six-arm DNA branched junctions. *Biochemistry* **30**, 5667–5674 (1991).
- Yan, H., Park, S. H., Finkelstein, G., Reif, J. H. & LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301**, 1882–1884 (2003).
- He, Y., Tian, Y., Ribbe, A. E. & Mao, C. D. Highly connected two-dimensional crystals of DNA six-point-stars. *J. Am. Chem. Soc.* **128**, 15978–15979 (2006).
- Hong, F. et al. Layered-crossover tiles with precisely tunable angles for 2D and 3D DNA crystal engineering. *J. Am. Chem. Soc.* **140**, 14670–14676 (2018).
- He, Y., Chen, Y., Liu, H., Ribbe, A. E. & Mao, C. Self-assembly of hexagonal DNA two-dimensional (2D) arrays. *J. Am. Chem. Soc.* **127**, 12202–12203 (2005).
- Kallenbach, N. R., Ma, R. I. & Seeman, N. C. An immobile nucleic-acid junction constructed from oligonucleotides. *Nature* **305**, 829–831 (1983).
- Seeman, N. C. & Kallenbach, N. R. Design of immobile nucleic-acid junctions. *Biophys. J.* **44**, 201–209 (1983).
- Srinivasan, A. R. & Olson, W. K. Computer models of DNA four-way junctions. *Biochem. (Easton)* **33**, 9389–9404 (1994).
- von Kitzing, E., Lilley, D. M. J. & Diekmann, S. The stereochemistry of a four-way DNA junction: a theoretical study. *Nucleic Acids Res.* **18**, 2671–2683 (1990).
- Wood, A. A., Nunn, C. M., Trent, J. O. & Neidle, S. Sequence-dependent crossed helix packing in the crystal structure of a B-DNA decamer yields a detailed model for the Holliday junction. *J. Mol. Biol.* **269**, 827–841 (1997).
- Goodsell, D. S., Grzeskowiak, K. & Dickerson, R. E. Crystal Structure of C-T-C-T-C-G-A-G-A-G. Implications for the Structure of the Holliday junction. *Biochem. (Easton)* **34**, 1022–1029 (1995).
- Litke, J. L., Li, Y., Nocka, L. M. & Mukerji, I. Probing the ion binding site in a DNA Holliday junction using Förster resonance energy transfer (FRET). *Int. J. Mol. Sci.* **17**, <https://doi.org/10.3390/ijms17030366> (2016).
- Várnai, P. & Timsit, Y. Differential stability of DNA crossovers in solution mediated by divalent cations. *Nucleic Acids Res.* **38**, 4163–4172 (2010).
- Duckett, D. R., Murchie, A. I. & Lilley, D. M. The role of metal ions in the conformation of the four-way DNA junction. *EMBO J.* **9**, 583–590 (1990).
- Møllegaard, N. E., Murchie, A. I., Lilley, D. M. & Nielsen, P. E. Uranyl photoprobing of a four-way DNA junction: evidence for specific metal ion binding. *EMBO J.* **13**, 1508–1513 (1994).
- Lu, X.-J., Bussemaker, H. J. & Olson, W. K. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.* **43**, e142–e142 (2015).
- Wang, W. et al. Holliday junction thermodynamics and structure: coarse-grained simulations and experiments. *Sci. Rep.* **6**, 22863 (2016).
- Yadav, R. K. & Yadava, U. Molecular dynamics simulation of hydrated d(CGGGTACCCG)4 as a four-way DNA Holliday junction and comparison with the crystallographic structure. *Mol. Simul.* **42**, 25–30 (2016).
- Yoo, J. & Aksimentiev, A. In situ structure and dynamics of DNA Origami determined through molecular dynamics simulations. *Proc. Natl Acad. Sci. USA* **110**, 20099–20104 (2013).
- Watson, J., Hays, F. A. & Ho, P. S. Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Res.* **32**, 3017–3027 (2004).
- Ledvina, P. S., Yao, N., Choudhary, A. & Quiocho, F. A. Negative electrostatic surface potential of protein sites specific for anionic ligands. *Proc. Natl Acad. Sci. USA* **93**, 6786 (1996).
- Auffinger, P., Bielecki, L. & Westhof, E. Anion binding to nucleic acids. *Structure* **12**, 379–388 (2004).



59. Trausch, J. J., Marcano-Velázquez, J. G., Matyjasik, M. M. & Batey, R. T. Metal ion-mediated nucleobase recognition by the ZTP riboswitch. *Chem. Biol.* **22**, 829–837 (2015).
60. Valls, N., Richter, M. & Subirana, J. A. Structure of a DNA duplex with all-AT base pairs. *Acta Crystallogr. Sect. D.* **61**, 1587–1593 (2005).
61. Šponer, J. et al. RNA structural dynamics as captured by molecular simulations: a comprehensive overview. *Chem. Rev.* **118**, 4177–4338 (2018).
62. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr., Pt A* **276**, 307–326 (1997).
63. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
64. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **58**, 1948–1954 (2002).
65. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D-Struct. Biol.* **53**, 240–255 (1997).
66. Bailey, S. The CCP4 suite - programs for protein crystallography. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **50**, 760–763 (1994).
67. Schrödinger, L. & DeLano, W. PyMOL (2020).
68. Case, D. A. et al. AMBER 18 (University of California, San Francisco, 2018).
69. Zgarbová, M. et al. Refinement of the sugar–phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA. *J. Chem. Theory Comput.* **11**, 5723–5736 (2015).
70. Ivani, I. et al. Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016).
71. Liebl, K. & Zacharias, M. Tumucl: a new accurate DNA force field consistent with high-level quantum chemistry. *J. Chem. Theory Comput.* **17**, 7096–7105 (2021).
72. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
73. Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
74. Le Grand, S., Götz, A. W. & Walker, R. C. SPFP: speed without compromise—a mixed precision model for gpu accelerated molecular dynamics simulations. *Comput. Phys. Commun.* **184**, 374–380 (2013).
75. Krepl, M. et al. An intricate balance of hydrogen bonding, ion atmosphere and dynamics facilitates a seamless uracil to cytosine substitution in the U-turn of the neomycin-sensing riboswitch. *Nucleic Acids Res.* **46**, 6528–6543 (2018).
76. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
77. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

## Acknowledgements

Results shown in this report are derived from work performed at the Argonne Photon Source (APS), Advanced Light Source (ALS), and the National Synchrotron Light Source II (NSLS-II). The ANL Structural Biology Center (SBC) at the Advanced Photon Source (SBC-CAT) is operated by UChicago Argonne, LLC, for the U.S. Department of Energy (DOE), Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The Berkeley Center for Structural Biology is supported in part by the Howard Hughes Medical Institute. The ALS is a DOE Office of Science User Facility under Contract No. DE-AC02-05CH11231. The ALS-ENABLE beamlines are supported in part by the NIH, National Institute of General Medical Sciences (NIGMS), grant P30 GM124169. Results from beamlines AMX (17-ID) and FMX (17-BM) at the NSLS-II,

which is a DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The Life Science Biomedical Technology Research resource is primarily supported by the NIH, NIGMS through a Biomedical Technology Research Resource P41 grant (P41GM111244), National Science Foundation Division of Materials Research (NSF2004250), and by the DOE Office of Biological and Environmental Research (KP1605010). This work was supported in part with projects SYMBIT reg. number CZ.02.1.01/0.0/0.0/15\_003/0000477 financed by the ERDF (M.K. and J.S.) and 21-23718S by the Czech Science Foundation (M.K. and J.S.). N.S. acknowledges startup funds from Arizona State University. H.Y., N.S., and P.S. gratefully acknowledge support from the National Science Foundation Division of Materials Research (NSF2004250). H.Y. was additionally supported by the Presidential Strategic Initiative Fund from Arizona State University.

## Author contributions

C.R.S. and H.Y. conceived of the project. T.M., A.B., and I.F. purified all DNA and prepared all crystals used in the work under the direction of C.R.S. and N.S., and C.R.S. performed all crystallographic data analysis and structure solution and refinement. M.K., J.S., and P.S. conceived of all molecular dynamics simulation experiments, and the analysis of the simulation data was performed by M.K. The experimental analysis of all junction angles was performed by M.M. under the direction of P.S. All authors discussed the results and provided manuscript feedback. C.R.S. and T.M. prepared all figures, and C.R.S., M.K., N.S., and H.Y. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30779-6>.

**Correspondence** and requests for materials should be addressed to Nicholas Stephanopoulos or Hao Yan.

**Peer review information** *Nature Communications* thanks Vlasios Mavrantzas and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022