COMPUTATIONAL
ANDSTRUCTURAL
BIOTECHNOLOGY
J O U R N A L

Review

In silico modelling of DNA nanostructures

Tadija Kekic, Ivan Barisic *

Austrian Institute of Technology GmbH, Center for Health & Bioresources, Molecular Diagnostics, Giefinggasse 4, 1210 Vienna, Austria

A B S T R A C T

The rise of material science and nanotechnology created a demand for a next generation of materials and procedures that can transcend the shaping of simple geometrical nano-objects. As a legacy of the technological progress made in the Human Genome Project, DNA was identified as a possible candidate. The low production costs of custom-made DNA molecules and the possibilities concerning the structural manipulation triggered significant advances in the field of DNA nanotechnology in the last decade. To facilitate the development of new DNA nanostructures and provide users an insight in less intuitive complexities and physical properties of the DNA folding, several *in silico* modelling tools were published. Here, we summarize the main characteristics of these specialised tools, describe the most common design principles and discuss tools and strategies used to predict the properties of DNA nanostructures.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creative-commons.org/licenses/by-nc-nd/4.0/).

Contents

* Corresponding author.
  E-mail address: Ivan.Barisic@ait.ac.at (I. Barisic).

1. Introduction

The last two decades of nanotechnology marked a pursuit for materials with reliable and modular properties that can be used as standards for the development of nanostructures. Out of a range of organic and inorganic candidates [1–7], the deoxyribonucleic acid (DNA) was recognised for its outstanding features. The biological role of DNA, as a carrier of information for the synthesis and regulation of virtually all biogenic elements, makes it a highly modular molecule. The unique predictability of interactions, that DNA tends to form, allows for the generalisation of its binding patterns. The most familiar are the Watson-Crick base pairing, Holliday junctions and quadruplexes. The reliability of these interactions dictates and modulates its hetero-dimerisation with extreme precision. Accompanied by high chemical stability, a low price and the unproblematic availability of custom oligonucleotides, the modularity of DNA led to the development of a variety of DNA-based nanostructures and nanomaterials of defined shapes [8–14].

Today, hundreds of DNA oligonucleotides can be routinely designed and synthesised to precisely self-assemble into stable nanostructures by creating tens of thousands of predefined interactions on the angstrom scale. The achieved complexity of these structures defines the DNA as one of the most versatile materials and a potential future standard material in nanotechnology. The development of DNA nanostructures and the organisation of the involved oligonucleotides are facilitated by *in silico* modelling tools [15–19]. In this review, we summarised the main characteristics of selected modelling and structural validation tools for DNA nanostructures. Most of these tools are based on one of four principal design strategies of DNA nanostructures (Table 1).

1.1. Tile-based DNA structures

In attempting to create new tools for structural biology, Ned Seeman proposed the development of the "immobile junction" DNA structure [20]. The structure was based on tightly packed, sequence-mediated interactions between distinct regions of several oligonucleotides. In the following years, the method was expanded into the tile-based DNA nanostructure design concept and was implemented for the construction of higher-order periodic and aperiodic lattices and simple 3D objects (Fig. 1) [21–24].

Before the emergence of the DNA origami technology, the tile-based DNA architecture was commonly used. In this first period, it was characterized by the limitations regarding oligonucleotide fabrication and *in silico* design support for more complex objects. This was changed by recently published state-of-the-art modelling tools that implement a special type of tile-based architecture, named DNA bricks, to assemble structures of up to 1 GDa in size [23–26]. These are made from tens of thousands of unique components. This advancement brought new possibilities to this DNA scaffold-independent technology.

1.2. Multilayer DNA origami

Developed by Paul Rothemund in 2006, the DNA origami approach of folding DNA is based on the hybridisation between two distinct types of single strand DNA, named 'staple' and

Table 1

Overall comparison of tools for the modelling of DNA nanostructures. In section "Architecture", M.DNA origami stands for Multilayer DNA origami, and W. DNA origami stands for Wireframe DNA origami. Tools can be accessed on following websites: Tiamat (yanlab.asu.edu); CaDNAno (cadnano.org); vHelix (vhelix.net); Daedalus (daedalus-dna-origami. org); Perdix (perdix-dna-origami.org); Adenita (samson-connect.net).

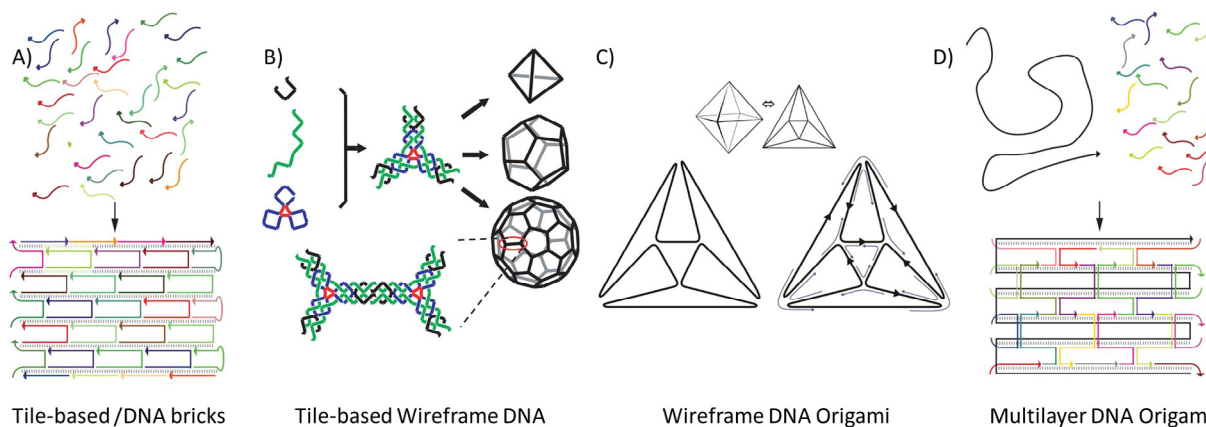| Software	Architecture	3D structures	Free-hand design	Sequence Editing	Sequence Optimisation	Structural Validation	Graphical interface	Large Designs	Mesh Input	Scalable Visualisation
Tiamat	Tile-based	+	+	+	+	–	+	–	–	+
CaDNAno	M. DNA origami	+	–	+	+	–	+	+	–	+
vHelix	W. DNA origami	+	–	+	–	–	–	+	+	–
Daedalus	W. DNA origami	+	–	–	–	–	–	+	+	–
Perdix	W. DNA origami	–	+	+	–	–	–	+	+	–
Adenita	M./W. DNA origami/ tile-based	+	+	+	+	+	+	+	+	+

Fig. 1. Principal design strategies for DNA nanostructures. Assembled from hundreds of short oligonucleotides that interact with each other in a predefined manner, tile-based DNA architectures can be used to form both (A) multilayer and (B) wireframe structures. (C, D) Alternative approaches are based on nanostructures comprising a long DNA scaffold that is folded by short staple strands. They are more stable with a better assembly yield, but are limited in size and complexity due to the scaffold's length and currently implemented routing algorithms; (A, D) Reprint from [21] with permission from AAAS; B) Reprint from [22] with permission from Springer Nature, copyright (2008); C) Reprint from [19] with permission from Springer Nature: Nature.
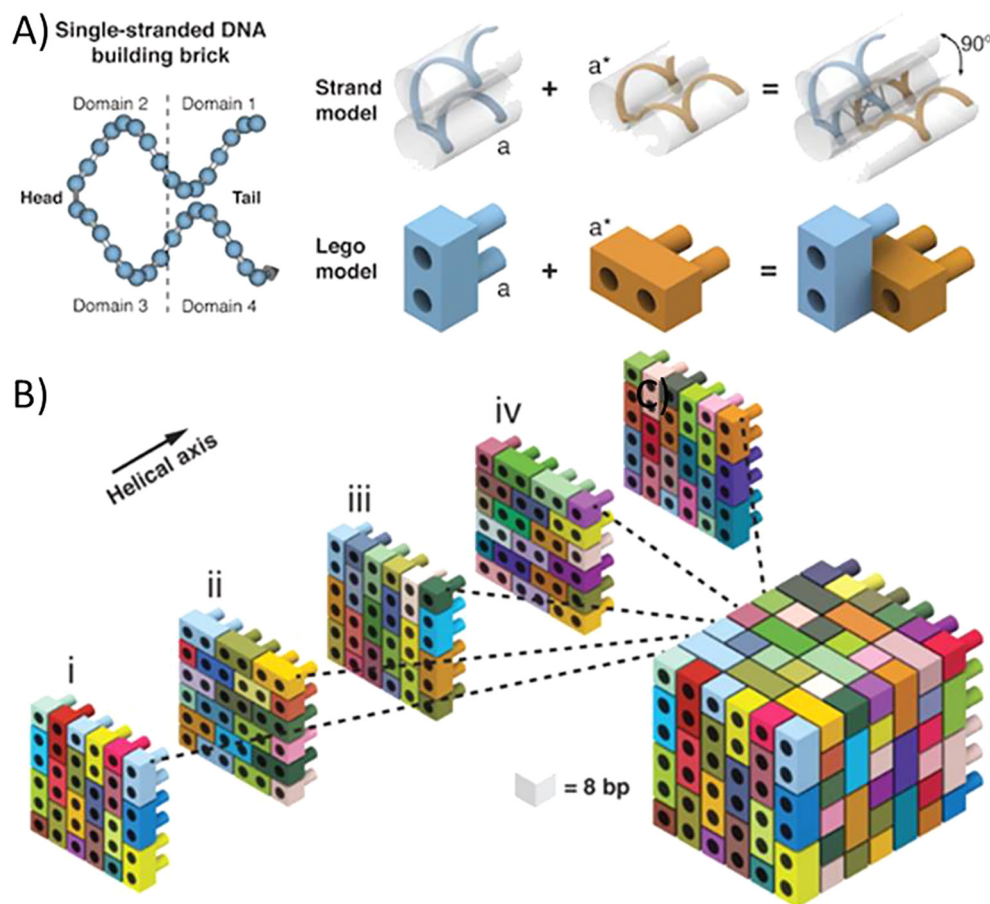
Tile-based /DNA bricks   Tile-based Wireframe DNA   Wireframe DNA Origami   Multilayer DNA Origami

Fig. 2. Assembly of DNA bricks into gigadalton structures. (A) Based on the principle of tile-based nanostructures, DNA bricks have four binding domains, each complementary to another predefined DNA brick. The local assembly can be compared to Lego bricks; (B) Due to a highly ordered organisation, this type of DNA nanostructures allows the development of consistent units and domains in larger, more complex assemblies. In contrast to scaffolded design concepts, various shapes of structures can be realised by changing a relatively small percentage of oligonucleotides from the total assembly pool. Reprint from [23] with permission from AAAS.

'scaffold' DNA strands [27]. The folding process is guided by the hybridisation of the custom-made sequence regions on the short staple oligonucleotides (20–60 bp) to their complementary regions on the far longer scaffold (ca. 7 kbp). As these complementary regions on the scaffold are usually distant from each other, the hybridisation between DNA molecules via crossovers brings them together (Fig. 1) [25,26]. By association of hundreds of different staple strands with a single scaffold, the user can precisely modulate their molecular organisation, and thus, use them for the development of nanostructures [27,28]. For more advanced implementations, it is possible to use chemically modified staple strand oligonucleotides allowing interactions with different nanostructures, proteins, metal and glass surfaces, and lipid membranes [8,12,28–33].

The nanostructures made by the multilayer DNA origami approach tent to be condensed and bulky and require high concentrations of Mg^{2+} or other positively-charged ions to remain stable in solutions [13,34]. For smaller constructs, the scaffold strand is usually obtained from M13 bacteriophages and can be assembled in a one-pot reaction. However, larger and more complex structures are assembled from several different scaffold strands. The experimental procedure is thus more complicated, as each part of the super-structure must be assembled separately and purified before the super-assembly begins. In comparison to other design strategies, the DNA origami nanostructures are rigid and highly visible by electron microscopy methods such as TEM and Cryo-EM [35–36].

1.3. Wireframe DNA origami and wireframe DNA

As a special design type of DNA origami, wireframe DNA origami structures allow almost arbitrary 2D and 3D geometries [37,38]. The main distinction of this concept is that it organises elements of the nanostructure to populate only the areas surrounding the edges of abstract geometrical objects. Advanced algorithms are required to calculate the optimised routing of the scaffold strand throughout this type of structure (Fig. 1) [17,19].

Although similar regarding the final DNA wireframe structures, the wireframe DNA method is a subtype of the tile-based DNA architecture, and as such, routing calculations are far simpler, and the maximum size of the structures can be significantly higher (Fig. 1) [19,39]. Both wireframe methods are used to create open and hollow nanostructures, usually more flexible in solution. In contrast to multilayer DNA origamis, they can be folded and are stable also under low concentrations of Mg^{2+} and physiological ion conditions [40].

1.4. DNA bricks

As proposed by Peng Yin in 2012, the method relies on the assembly of thousands of short oligonucleotides, named DNA bricks [23]. Each DNA brick comprises four short domains, capable of binding four other bricks (Fig. 2). This allows the construction of arbitrary shaped nanostructures [23,26]. The DNA bricks are 32 or
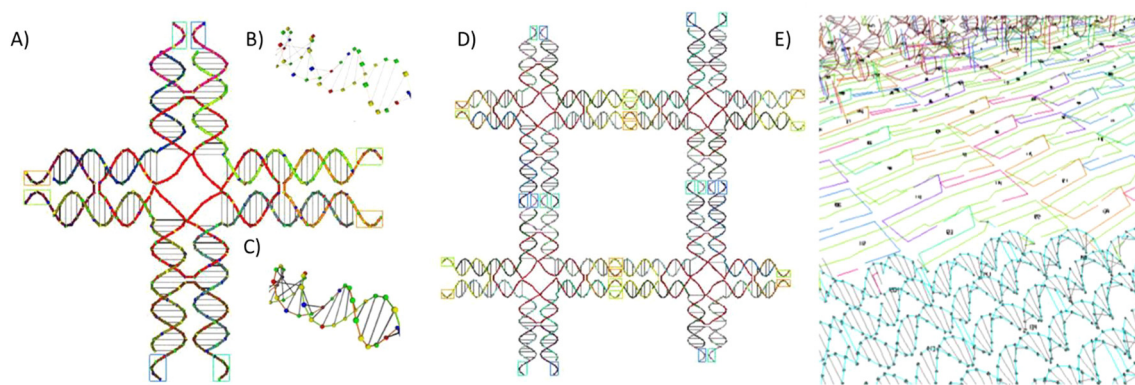
Fig. 3. Design process and visualization in Tiamat. (A) Representation of a 4x4 DNA tile created in the Tiamat. Artefacts are visible on the vertices. The models can be visualized in a (B) simplified and (C) rendered mode. (D) A model of a larger, planar structure developed by the integration of several tile-based segments. (E) A scalable visualization of the structures can be created manually. Adapted from [15].

52 bp in length. The stability and total size of this type of nanostructures are strongly dependent on the length of individual brick's binding domains [23,26]. In contrast to DNA origamis, bricks-made objects are more flexible due to the lower stability of their (single) junctions, and its more opened organisation exhibits significantly higher ion permeability [41]. More importantly, the design strategy is not limited by the length of a scaffold strand. As such, this type of architecture can be used for the construction of significantly larger structures, often created in one-step assembly procedures (Fig. 2) [26]. To simplify the design of DNA brick nanostructures, the commercially available modelling tool Nanobricks was developed in 2012 [26].

2. Software solutions for DNA-based modelling

The first software solutions for DNA nanostructure design were based on existing molecular modelling tools using standardised structural files (e.g. pdb, mol, mae) [42–44]. These tools described or introduced modifications to the DNA structure on an atomic level. In these general molecular modelling tools, the level of details was optimised for a work with proteins and smaller molecules. Modelling of a large-scale DNA object, however, required a drastic increase in computational power, making these programs unusable in desktop environments. Due to this, early DNA nanostructure's modelling was done mostly manually on the level of the sequence, or by using command line driven programs and basic 3D renderings of low-complexity constructs [15]. The lack of structural details in the design process limited the ability to design and manipulate multi-layered 3D structures. This, accompanied by factors such as ease of fabrication and better structural imaging possibilities, led to the fact that mostly two-dimensional, plane DNA nanostructures were presented in early studies. Also, 3D DNA nanostructures were developed, however, they required a significant amount of manual calculations and proficiency in the use of general modelling tools [15]. Driven by the functionalisation requirements and ever-increasing complexity of the DNA nanostructure designs, several specialised tools for their modelling were developed in the last decade [19,26,15–17,28–47].

First, specialised, top-down software solutions were focusing on major optimisations and adaptations of the existing tools to better fit the geometry of the DNA molecule [16]. The high computational requirements were diminished by making a single nucleotide base the smallest building block of the model (Fig. 3) [15–17]. Strong positional and orientational restrictions were introduced to preserve a relatively high flexibility during the design process. The architecture of planar nanostructures is relatively unrestrained

[15]. However, the DNA origami architectures are forced in either square or honeycomb lattice organisations [16]. These organisational types, characterised by the multitude of linear DNA fragments of similar length and placed in parallel, remained dominant to this day. The parallel placement of DNA fragments and the predictable helical turns were crucial in the optimisation of bridging points and crossovers by which staple strands force the proximity between distant areas of the scaffold DNA. This allowed also the development of more intuitive modelling solutions and sophisticated validation methods [18]. In the following sections, we will discuss freely available *in silico* tools for the modelling and analysis of DNA nanostructures.

2.1. Tiamat (2008)

The open-source software Tiamat is an early-stage first-generation modelling tool specialised for DNA nanostructures. It addresses two main limitations of the general molecular *in silico* tools [15]. First, Tiamat integrates sequence generation and manipulation. Second, in terms of appropriate visualisation and graphical requirements, the software was developed for complex planar structures. By a strong focus on the base-pairing aspect of the modelling, Tiamat offers, to an experienced user, the ability to design 3D structures with a relatively high freedom, unrestrained by lattice-type architectures [15]. The trade-offs to such a degree of freedom are a low theoretical value of the model and a lack of *in silico* validation options of the nanostructure (Fig. 3). Following the design process, the DNA sequences are analysed to prevent dimerisation artefacts, while considering the constraints set for the sequence length, repetition limits, GC content and melting temperatures [15]. Due to a minimalistic graphical user interface, the modelling process of high-complexity designs is extremely time-consuming. However, the problem that the rendering of larger nanostructures creates considerable computational demands is well solved by Tiamat. To reduce visualisation requirements, Tiamat allows the use of proxy geometries (e.g. straight lines instead of helices) which can be reversed based on the user's preferences (Fig. 3) [15].

2.2. CaDNAno (2009)

The open-source software package CaDNAno was published as a modelling tool to enable computational support for multilayer DNA origami structures. In the first version, only a honeycomb lattice was available (Fig. 4) [16]. With this architecture type, each DNA helix has up to three neighbouring helices. In a canonical CaD-
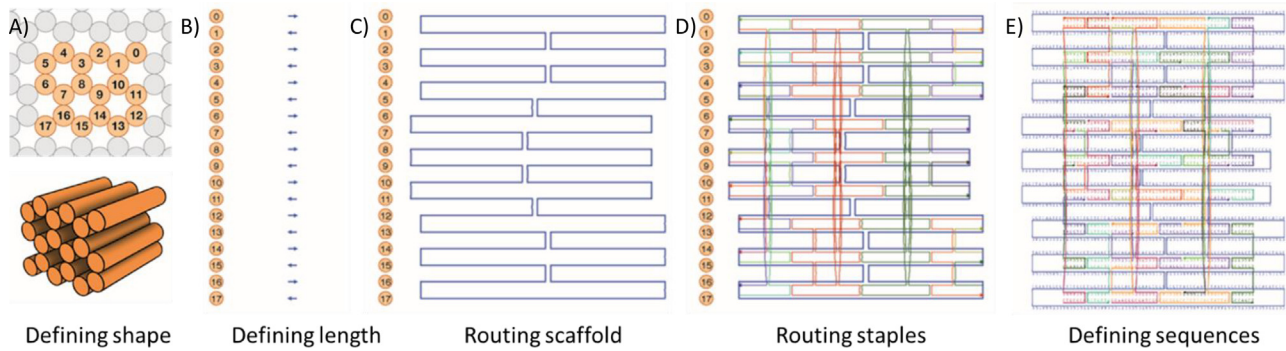
Fig. 4. Design process in caDNAno. (A) Multilayer DNA origamis can be modeled in CaDNAno by defining the lattice organization of the structure on the 2D, top-down scheme. (B, C) The 3D shape and more detailed organization of the nanostructure is set manually in the diagram viewer by defining the length of the individual DNA helices. As the interactions between different regions of the scaffold are represented in the two-dimensional diagram, the user can confirm the relaxed shape of the model in external, structure validation tools. Adjusted from [16]. Made available under the terms of the Creative Commons Attribution-NonCommercial Licence 2.0 (CC BY-NC).
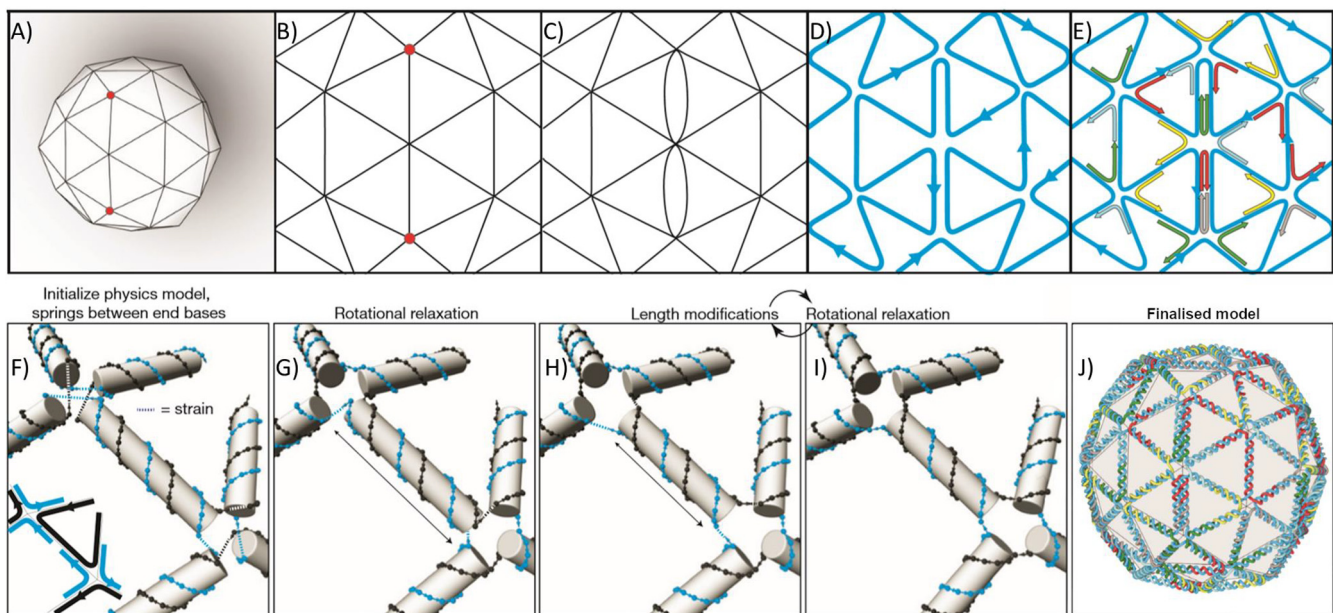

Fig. 5. Routing and relaxation in vHelix. (A) Through an automatized process the user-defined mesh of the desired DNA nanostructure is triangularized and adjusted for (B–D) scaffold routing by the A-trails method. This is followed by (E) routing the staple strands and (F–I) a relaxation process resulting in the structural approximation of the nanostructure to the initial mesh design (J). Reprint from [19] with permission from Springer Nature: Nature.

NAno model, antiparallel crossovers between neighbouring helices occur every 21 bases at positions where the involved nucleotides have the closest distance to each other. With three adjected helices, this would translate to a maximum of one stable crossover every seven bases [16]. CaDNAno allows for a strong deviation of the canonical crossover rule, often required for the design of a complex nanostructure [10,12,16]. However, integrated tools such as the automatic design of staple strands tend to create artefacts once the design rules are violated. Significant improvements to the usability of the software were done by the introduction of the square lattice design option and a simple 3D visualisation plugin for Maya. CaDNAno can export models in the .json file format, DNA sequences in the .csv format, and schematic views in the .svg format.

The modelling process of CaDNAno can be summarised in six main steps (Fig. 4). First, the planar shape of the nanostructure is defined by populating the empty positions on the lattice. Second, the length of the individual helices is defined. Third, a DNA-origami scaffold strand is routed through the 2-D diagram of the nanostructure. Fourth, the staple strands are calculated and routed through the structure. Five, the length and routing of the staple strands is adjusted to the user needs. Six, the DNA sequences of the scaffold and other elements of the structure are defined by the choice of the appropriate scaffold plasmid [16].

2.3. vHelix (2015)

The modelling of wireframe DNA origami nanostructures can be facilitated by the open-source tool vHelix. The design process starts in an external program by defining a polyhedral mesh that is used as an input (Fig. 5). The edges and vertices of the mesh serve as a representation of the geometry, organisation and interactions that the DNA forms in the structure [19].

The initial step in the mesh processing is its triangularisation with the purpose of optimising structural stability. This is followed by the implementation of the graph theory in solving the routing of the scaffold strand throughout the structure (Fig. 5) [19]. In the implemented routing method using A-trails (specific type of Eule-
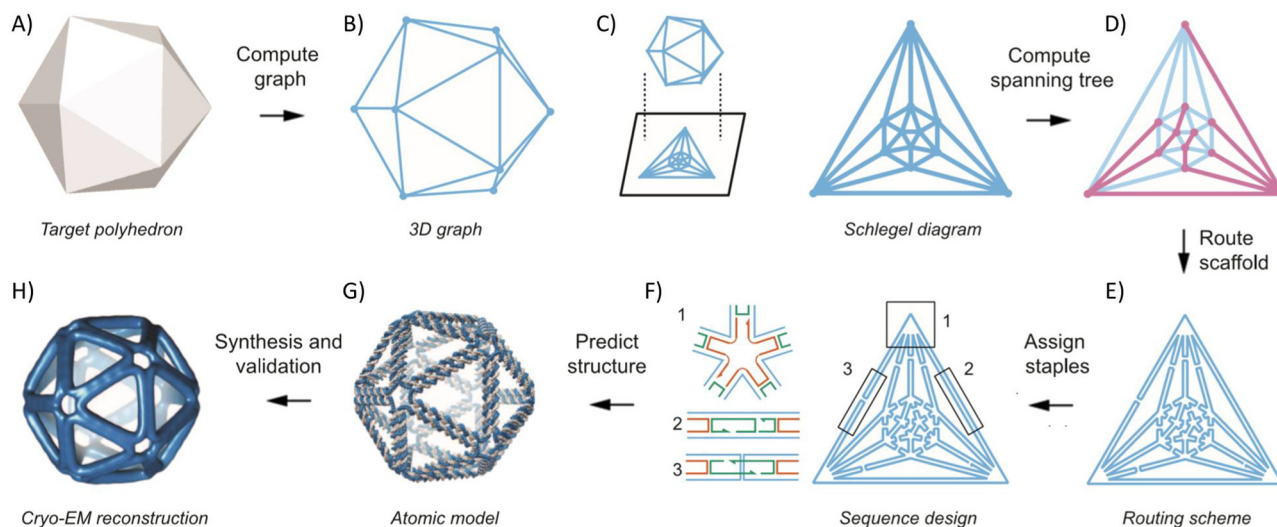
Fig. 6. Automated routing of a user-defined mesh in Daedalus. Following the (A) mesh design, a fully automatized procedure computes (B) a 3D graph and (C-D) a spanning tree. The spanning tree is used in (E) routing the scaffold, followed by (F) the placement of staple strands. The resulting structure can be exported at an (G) all-atom level and used for further validations by computational or experimental methods. Reprint from [17] with permission from AAAS.

rian circuits), consecutive edges in the circuit are always neighbours in the cyclic ordering around the vertices. The result of this method is that the scaffold DNA passes in most cases one to three times through all defined edges of the mesh, and the staple strands are routed implicitly from the scaffold [19]. Although the A-trail routing is not the most efficient algorithm, vHelix calculates it in seconds. This is achieved by the implementation of a systematic search and use of pruning and heuristics in the branching [19].

To reduce artefacts of the routing procedure, a physical model is used to relax and distribute molecular strain throughout iterations of rotational and longitudinal relaxations of the DNA helices (Fig. 5). In the relaxation process, each helix is treated as a stiff rod with weighted connections between bridging nucleotides [19]. Following the relaxation, the vHelix model can be imported into Autodesk Maya for further visualisation and manual post-processing. As non-restricted, non-lattice modelling of 3D structures often leads to a strong bending between neighbouring edges, vHelix allows the introduction of additional, unpaired nucleotides to relive the strain on the affected areas of the structure [19]. vHelix is best suited to design flexible and open wireframe DNA origami nanostructures with very long scaffold strands. The software focuses on these structures as they are expected to be more stable in low salt conditions than compact multilayer DNA origamis or tile-based DNA structures.

2.4. Daedalus (2016)

Similar to vHelix, the open-source software Daedalus offers a fully automated design of robust wireframe DNA assemblies based on an initial polyhedral mesh as input (Fig. 6). It has no graphical user interface and requires Matlab. A main difference to vHelix is that Daedalus does not perform a triangulation of the mesh. The designed object is fully defined through spatial coordinates of all vertices, the connectivity of edges between vertices and the faces to which vertices belong [17].

To ensure an optimised routing with the Eulerian circuit method, the sequence length of all edges must be multiples of 10.5 bp. The routing procedure is initiated by a minimum weight spanning tree generated with the Prim's algorithm (Fig. 6). A single scaffold crossover is placed in the centres of edges that are not part of the spanning tree. Based on both crossover positions and the spanning tree, the rest of the routing is calculated so the scaffold does not interact with himself at the vertices. Due to the routing algorithm, the scaffold passes through every edge twice [17].
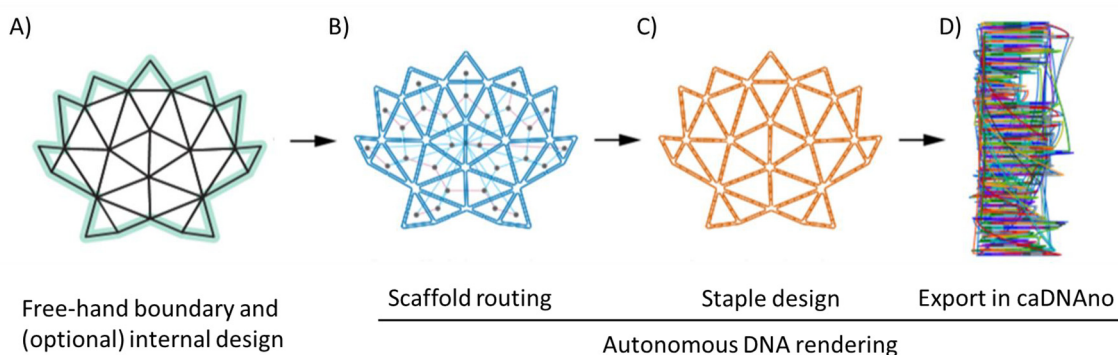

Fig. 7. Scheme of the free-hand mesh design and routing process in PERDIX. (A) An arbitrary two-dimensional mesh structure, with or without internal organization, is inputted into the tool. (B, C) The autonomous, dual-graph routing of the scaffold and staple strands is achieved by converting each edge into parallel duplexes joined by antiparallel crossovers. Following the routing, the model is exported in a caDNAno readable format for external processing (D). Adjusted from [46]. Made available under the terms of the Creative Commons Attribution-NonCommercial Licence 4.0 (CC BY-NC).
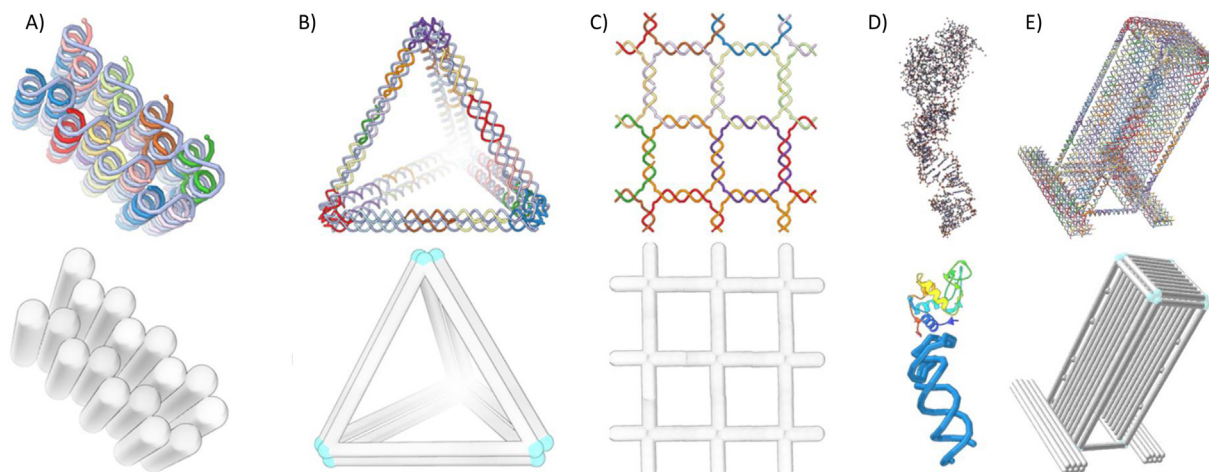
Fig. 8. Design examples and visualizations that can be realized in Adenita. The tool is compatible with (A) multilayer DNA origamis, (B) Deadalus wireframe structures, (C) tile-based structures, (D) proteins and aptamers, and (E) nanostructures comprising several different design concepts. Adapted from [47] with permission from the authors.

Following the scaffold routing, the staple strands are determined based on the Watson-Crick pair complementarity, and form either vertex staples or edge staples (Fig. 6). Vertex staples hybridise to the 10 or 11 nucleotides on the edges that are closest to the vertices, with a final length of 52 or 78 bp [17]. To optimise the folding, the algorithm will place major groves inwards at vertices. Edge staples hybridise within the central area of the edges and stabilise the structure by creating crossovers every 10, 11 or 21 bp, respectively. As major benefits of this routing algorithm, authors state the uniqueness of each spanning tree [17]. This allows designs with different routings for structures of the same shape. Daedalus provides an output in a simple mesh format, and a list of DNA sequences. Post-modelling refinements of the model on e.g. the atomic level must be conducted with external tools, which often results in the creation of artefacts due to the unpaired nucleotides of the vertices.

2.5. Perdix (2019)

Perdix is an autonomous modelling tool specialised for the development of arbitrary 2D objects from a user-defined 2D mesh (Fig. 7). As well as previously described tools that utilise a 3D mesh, the Perdix routing is based on the Eulerian circuit method. Perdix is not limited by discrete lengths of the edges. However, a minimum length of 38 bp with two double crossovers is obligatory [46]. This design freedom comes from the ability of the algorithm to introduce unpaired scaffold nucleotides and extend helices to the intersection of neighbouring helices making a new vertex (Fig. 7). After inputting a 2D mesh, the tool uses the DistMesh or Shapely algorithms to fill the interior of the mesh [46].

The routing process begins by translating all edges into two antiparallel scaffold lines. By connecting the endpoints of these lines over vertices, every edge becomes part of a small loop inside one larger loop structure. This is followed by a screening for all scaffold crossovers positions that can be formed between neighbouring loop pairs, finally resulting in a loop-crossover structure [46]. The Eulerian circuit method is then used to ensure that every vertex contains an even number of duplexes and that closed scaffold loops do not exist. Routing is done by the Prim's algorithm [46]. Following the routing, staple strands, 20–60 bp in length, are introduced (Fig. 7). On the unpaired area of the vertices, staple strands will create poly(T) loops. Also, circularised staple strands are properly nicked [46]. Finally, sequences are assigned to the scaffold and staple strands, and an atomic model is provided for an external post-processing. Perdix allows the integration of the

various functional molecules into the nanostructures, including molecular dyes, proteins and semiconductor nanocrystals. Perdix can export DNA origami structure in the caDNAno format (Fig. 7).

2.6. Adenita (2020)

Adenita was developed as an open-source plugin for SAMSON Connect, a 3D modelling graphical framework. Adenita attempts to reduce the modelling constraints of the earlier DNA-origami software solutions by allowing a user to work with all common types of DNA architecture simultaneously (Fig. 8).

Through implementation of a multiscale, adaptable visualisation of the model, Adenita simplifies the free-hand design of nanostructure's elements by adapting the structural complexity of the model to the needs of the current design state [47]. As such, large, robust structures can be built on a graphically abstract or mesh level, routings of scaffolds and staples can be represented and modified as tubes, and individual residues can be inspected and modified as all-atom models (Fig. 8). With this, Adenita provides a platform on which, an experienced user, can incorporate all existing design architectures into a single construct opening new opportunities to the field. Adenita facilitates also the integration of organic and inorganic molecules, including proteins, directly into the model of the DNA nanostructure (Fig. 8) [47]. This can significantly simplify the process to functionalise nanostructures, and the user's ability to investigate and modulate interactions between various functional elements, e.g. FRET pairs, protein binding regions or target-aptamer interactions.

Adenita is compatible with CaDNAno designs allowing an import of these structures. Furthermore, the Deadalus algorithm was implemented to create wireframe DNA origami structures. Adenita also offers a repository of parametrised structures that can be selected as a template or building block. Through the design process, an initial *in silico* validation of the structural properties can be performed by calculating the thermodynamic parameters by the nthal algorithm, from the Primer3 suite [47]. Advanced *in silico* validation can be done by exporting models in a run-ready, oxDNA format. This allows a simple transition from a modelling tool to a coarse-grained molecular dynamic software.

3. Validation of DNA nanostructure *in silico*

The parallel architecture of the multilayer DNA origami nanostructures is often not suitable for some sophisticated designs. Furthermore, the modelling of structures with higher degrees of
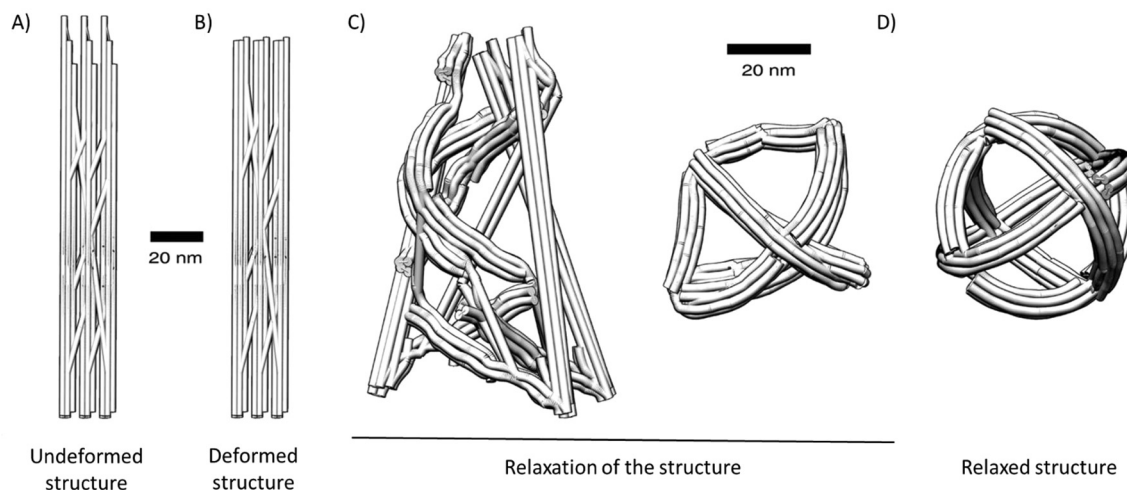
Fig. 9. The process of DNA nanostructures relaxation and the model validation in CanDo. (A, B) In the structural prediction process, the original, undeformed structure is deformed to reduce axial and torsional mismatch in neighboring helices, while their crossovers segments are fully constrained [18]. (C) The mechanical perturbation is done by twisting and turning individual helices until the shape of all crossovers is optimized. (D) This is followed by the minimization and a normal mode analysis to approximate the mechanical flexibility of the DNA nanostructure. Adjusted from [51]. Made available under the terms of the Creative Commons Attribution-NonCommercial Licence 3.0 (CC BY-NC).

structural freedom requires a validation of shape and stability. To solve these problems and reduce time-consuming *in vitro* evaluation experiments, several *in silico* structure validations solutions emerged in the recent years [18,48–50]. The structure validation solutions predict the dynamics of the simulated system by using force fields. The force fields mathematically describe interactions between elements of the system, which is required for molecular dynamics and mechanical energy perturbations. The precision of the force fields determines the precision and reliability of this interactions, and through it the level of details that can be investigated. For larger systems, use of highly precise, all-atom force fields (e.g. AMBER, CHARMM) leads to the drastic increase in computational requirements, as such, use of coarse-grained force fields (CanDo, OxDNA) is more convenient and time-conserving. In the following section, tools are ordered by the level of details they provide to a user – from lowest to highest (all-atom).

3.1. CanDo (2011)

CanDo is an online service for DNA origami nanostructures predicting the 3D shape and mechanical flexibility of the modelled structures based on their sequence connectivity maps. Beside the canonical twist, bend and stretch stiffness of the double-helical DNA domains, the prediction model considers nicks in the DNA helix, entropic elasticity of ssDNA and the effect of distant crossover structures [18]. In the first version, CanDo accepted only lattice-type models made in caDNAno. In following updates, the tool implemented also the use of non-lattice caDNAno designs and free-hand models made in Tiamat [51].

In CanDo, the designed structure is described as an association of isotropic elastic rods with finite element beams that model stretching, twisting and bending mechanics of the DNA (Fig. 9) [18]. The base pairs are described as finite element nodes with a defined position and orientation with rigid crossovers between them. The relaxation of the structure starts with the mechanical energy perturbation through which double helices deforms (via stretching and twisting) until the shape of all crossovers is optimised (Fig. 9) [18]. Forces allocated to the mechanical processes could be empirically determined by comparison with experimental TEM images and were consistent with the base-stacking free energy of DNA [18]. Following deformation, the nanostructure

requires the energy minimisation cycle to reach the relaxed conformation. The minimum of free energy is computed iteratively by non-linear finite element formulations, followed by the normal mode analysis. The final solution is determined by the Root-mean-square-fluctuations of 200 lowest normal modes through the equipartition theorem of statistical thermodynamics (Fig. 9) [18]. By providing a reliable way to ascertain the stability of the designed structure CanDo can greatly enhance the speed and quality of the design-test-design cycle.

3.2. oxDNA (2012)

As an open-source molecular dynamics coarse-grained simulation software, oxDNA allows user to investigate the thermodynamic properties of DNA-based nanostructures and to observe folding events that occur on the microsecond timescale [49]. DNA is defined as a string of nucleotides, where each nucleotide is a set of the rigid atom groups (deoxyribose sugar, phosphate group and nitrogenous base). Each nucleotide consists of several collinear interaction sites and a vector vertical to the plane of the base. (Fig. 10) [52]. The potential energy of the system is defined through hydrogen bonding, cross and coaxial stacking and stacking interactions, backbone potential and isotropic excluded volume interactions (Fig. 10) [48,49]. The parameters were selected based on their ability to reproduce melting temperatures as predicted by Santa Luciás nearest-neighbour model. For that reason, oxDNA ignores several important aspects of DNA modelling [48,49].

The oxDNA algorithm treats DNA as a symmetrical structure, and ignores major and minor groves, the variable size of the backbone sites and the structure nucleotides [48,49]. The model allows base pairing only between Watson-Crick pairs and makes no distinction between nucleotides in terms of their interaction strength [48]. OxDNA offers valuable insights into the local and global dynamics of the DNA nanostructure through the investigation of sequence dependent phenomena, such as stacking transitions, hybridisation free energy profiles, hairpin stability and dynamics of the single stranded DNA [48,52]. OxDNA can point out the structural abnormalities that rise from the cumulative effect of this phenomena allowing a fast and detailed structural validation of the designed model (Fig. 10).
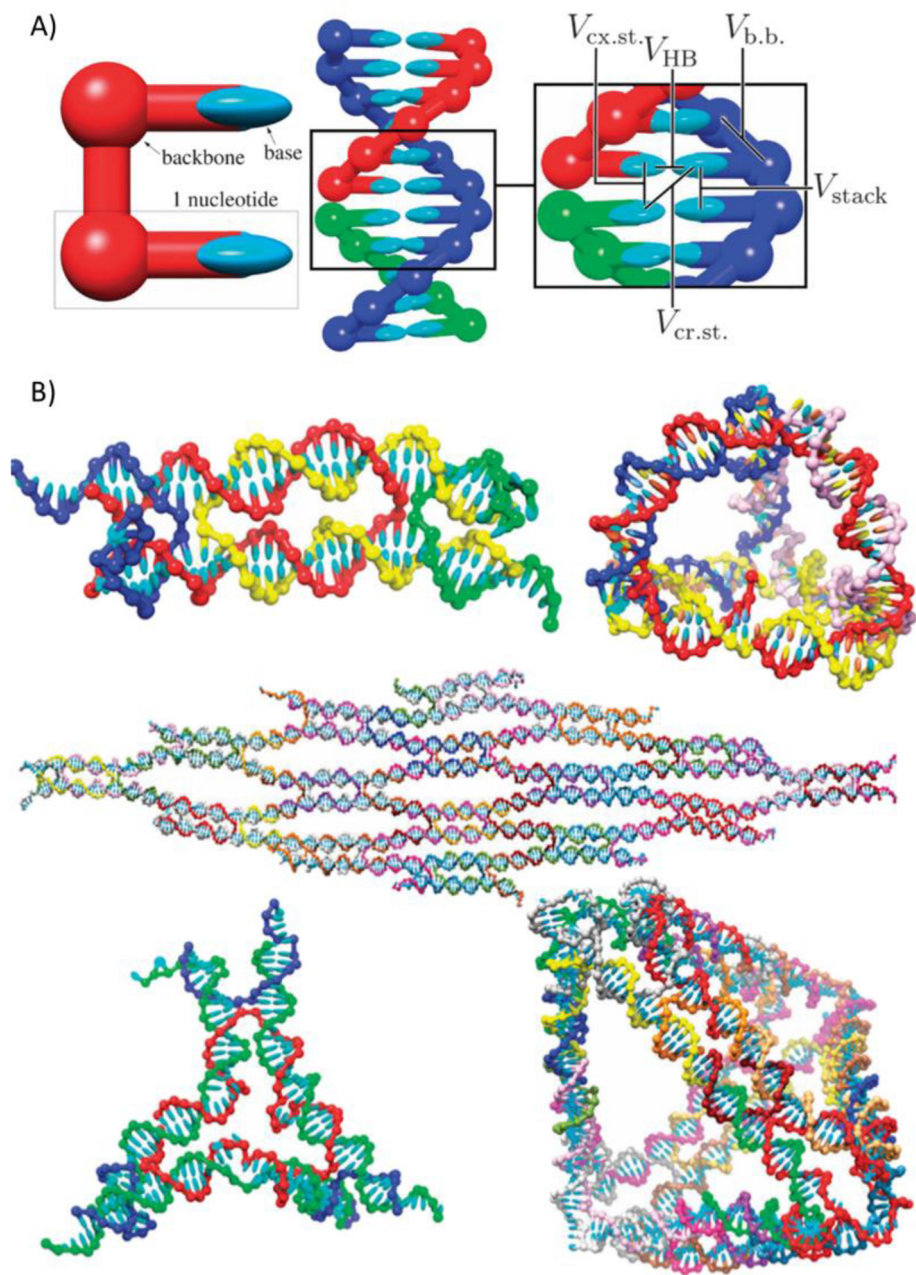
Fig. 10. Coarse-grained molecular dynamics in oxDNA. (A) Each nucleotide in the oxDNA model is a rigid body defined by a position and orientation of the backbone and base. Interactions between nucleotides are parameterized to best fit Santa Lucia's nearest-neighbor model. The DNA helix has a symmetrical structure without major and minor groves. (B) OxDNA can simulate molecular dynamics of DNA-based nanostructures on the microsecond timescale. This allows investigations of cumulative effects of various folding phenomena. As oxDNA is not built only for the analysis of DNA nanostructures, it accepts all types of DNA architectures. Reprint from (52) with permission from RSC Publishing.

3.3. Namd (2005)

The investigation of dynamic processes of DNA nanostructures on an all-atom level is computationally very expensive and requires specialised computer cluster. Further, it requires high-performance molecular dynamics (MD) software capable of making parallel calculations on hundreds or thousands of central and graphical processing units (CPU and GPU). As commonly used MD simulation tools such as AMBER and GROMACS are less suites for so high parallelisation, NAMD was developed [50]. NAMD is a scalable MD software capable of simulating all-atom systems of both DNA and proteins that comprise millions of atoms, both in

stable and varying conditions [50]. NAMD relies on AMBER and CHARMM parameters and retains general restrictions of all-atom MD such as very short simulation times.

In comparison to a coarse-grained model, NAMD is best used for the investigation of interactions between DNA nanostructures and membranes or integrated functional elements like proteins and other organic molecules (Fig. 11). To traverse easier over high energy barriers of the conformational landscape of the nanostructure, NAMD also offers steered and interactive molecular dynamics (SMD, IMD) [50]. In this type of MD, constant or harmonic forces are applied to a part of the structure with the purpose to force the transition to the other stable conformations of the system. As
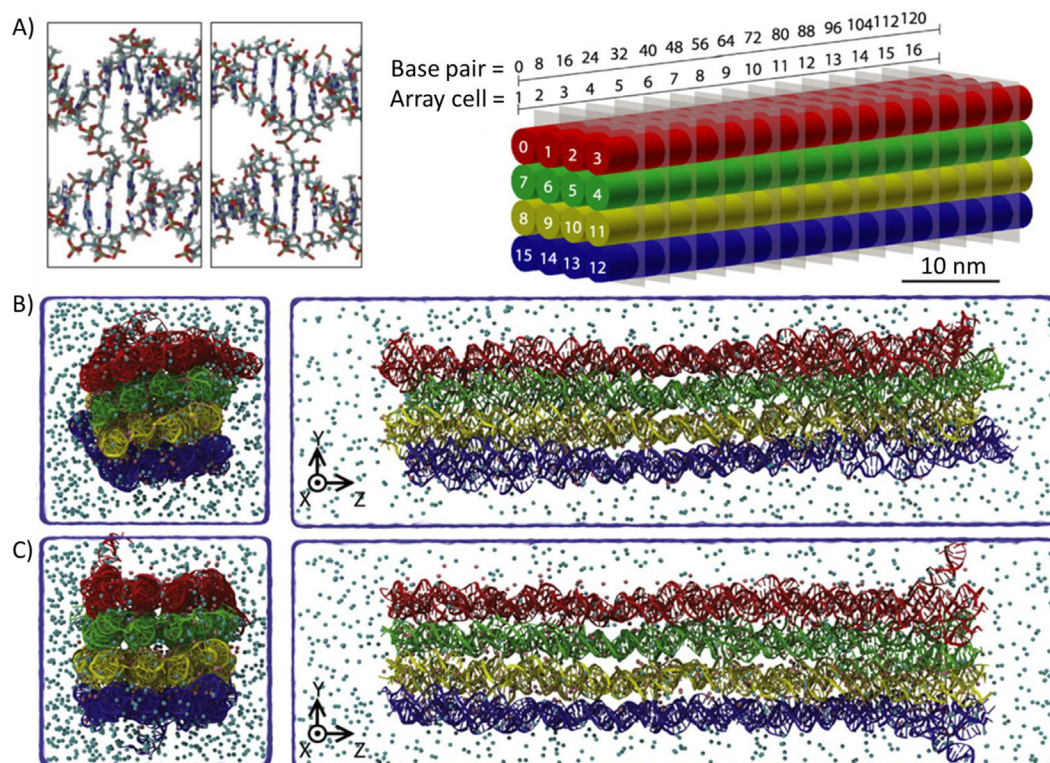
Fig. 11. Scalable all-atom molecular dynamics in NAMD. (A) Thermodynamic properties of DNA nanostructures can be investigated with NAMD. In their work, Slone et al (2016) compared the structural stability of idealized rod-shaped object built by (B) multilayer DNA origami and the (C) DNA bricks design strategies [42]. They noticed a significantly lower stability in the structure made from DNA bricks and attributed this to the difference in the density and type of junctions between the two design strategies. Adjusted from [41]. Made available under the terms of the Creative Commons Attribution-NonCommercial Licence 3.0 (CC BY-NC).

such, NAMD can be used to investigate mechanical properties of nanostructures in a way complementary to single molecule experimental techniques such as atomic force microscopy [50].

In 2016, Maffo et al. showcased the use of an elastic network-guided approach (excluding water molecules and ions) to reduce computational costs of DNA origami simulations by a factor of 10,000. The inter-helical electrostatic repulsion was treated through the addition of harmonic bonds with rest lengths of 31 Å between pairs of phosphorous atoms of every base pair [53]. Consequently, the staples strands could not dissociate from their scaffolds complement. Following the short elastic network-guided simulation, the obtained structure could be submerged in a solvent of explicit water molecules and run in a standard procedure [53]. By following the described procedure, the authors reported significantly lower RMSD values (referenced by a cryo-EM) in comparison to methods without the elastic network-guided step.

3.4. Post-modelling optimisation

Following modelling and structural validation steps, the oligonucleotide sequences used in the assembly of the nanostructure are exported from the *in silico* tools and can be individually inspected for unwanted behaviour like homo/hetero hybridisation, formation of structural artefacts, binding strength, melting temperature, and unspecific binding to a scaffold strand. The analysis of the oligonucleotides is a common procedure and a variety of the publicly available analysing tools exist [54,55]. It is important however, to consider assembly conditions, such as the concentration of the Mg^{2+} and Na^+ ions as well as the concentration of oligonucleotides in the calculation of these parameters.

The post-modelling optimisation of sequences is the last *in silico* step during the design of DNA nanostructures. As such, it is fol-

lowed by pooling the oligonucleotides in a buffer with sufficiently high Mg^{2+} concentration and assembly of the nanostructures in the temperature gradient [56,57].

4. Summary and outlook

Due to the increasing complexity of DNA nanotechnology, the development in the field is significantly influenced by the limitations regarding the available modelling and simulating software. To facilitate the implementation of these technologies, the ease-of-use and intuitive design in specialised *in silico* tools is very important. Although our ability to design and fabricate various types of DNA architecture has improved, the question of incorporating multiple architectures in the same nanostructure with one routing method still presents a significant challenge. Further, as almost no software allows the integration of molecules other than DNA or RNA, it is very hard or impossible for the common user to work with functionalised nanostructures *in silico*. Finally, improvements are required to efficiently visualise large, dynamic and 3D structures that cannot be represented by a simple lattice diagram [58,59]. This could be solved by exploiting the existing solutions implemented in industrial modelling tools. Along with the graphical optimisation, and ability to interact with the model on various levels of details, all indicated improvements are essential in exploiting the maximal potential of DNA as nanomaterial.

Strong progress has been done regarding the validation of the structural stability and investigation of the biophysical properties of DNA nanostructures. More advances will be beneficial regarding the simulating environment. Further, we lack the ability to efficiently simulate DNA-protein hybrid nanostructures with coarse-grained, or coarse-grained / all atom models. Improvements in the overall efficiency of both coarse-grained and all atom simula-

tions are necessary if methods are to be used without advanced computational infrastructure. Finally, the automatization of the analysis and optimisation of analysing tools specialised for the validation of DNA nanostructures can both simplify the design process and allow an accumulation of standardised data for various nanostructures. This would also further improve the currently used prediction models.

Author contributions

T.K. wrote the manuscript. I.B. reviewed and edited the manuscript. Authors declare no competing interests.

Funding

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wang J et al. High-efficiency and broadband near-infrared bi-functional metasurface based on rotary different-size silicon nanobricks. Nanomaterials 2019;9:1744.

[2] Jiang L et al. Tuning intermolecular noncovalent interactions for nanowires of organic semiconductors. Nanoscale 2010;2:2652–6.

[3] Zhao Y et al. Tunable emission from binary organic one-dimensional nanomaterials: an alternative approach to white-light emission. Adv Mater 2007;20:79–83.

[4] An BK et al. Color-tuned highly fluorescent organic nanowires, nanofabrics: easy massive fabrication and molecular structural origin. J Am Chem Soc 2009;131:3950–7.

[5] Busby M et al. Interactions of perylene bisimide in the one-dimensional channels of zeolite L. J Phys Chem C 2011;115:5974–88.

[6] Bottari G et al. Covalent noncovalent phthalocyanine−carbon nanostructure systems: synthesis, photoinduced electron transfer, and application to molecular photovoltaics. Chem Rev 2010;110:6768–816.

[7] Giner-Casares JJ et al. Inorganic nanoparticles for biomedicine: where materials scientists meet medical research. Mater Today 2016;19:19–28.

[8] Ketterer P. et al. Nanoscale rotary apparatus formed from tight-fitting 3D DNA component, Sci Adv 2016;2;e1501209.

[9] Castro CE et al. A primer to scaffolded DNA origami. Nat Methods 2011;8:221–9.

[10] Dietz H, Douglas SM, Shih WM. Folding DNA into Twisted and Curved Nanoscale Shapes. Science 2009;325:725–30.

[11] Sun S et al. DNA polygonal cavities with tunable shapes and sizes. Chem Commun 2015;51:16247–50.

[12] Doane TL, Alam R, Maye MM. Functionalization of quantum rods with oligonucleotides for programmable assembly with DNA origami. Nanoscale 2015;7:2883–8.

[13] Whitesides GM, Grzybowski B. Self-assembly at all scales. Science 2002;295:2418–21.

[14] Wagenbauer KF et al. Gigadalton-scale shape-programmable DNA assemblies. Nature 2017;552:78–83.

[15] Williams S. et al. Tiamat: A three-dimensional editing tool for complex DNA structures (2008) DNA.

[16] Douglas SM et al. Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic Acids Res 2009;37:5001–6.

[17] Veneziano R. et al. Designer nanoscale DNA assemblies programmed from the top down, Science 2016:352;1534.

[18] Kim D et al. Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructure. Nucleic Acids Res 2012;40:2862–8.

[19] Benson E et al. DNA rendering of polyhedral meshes at the nanoscale. Nature 2015;523:441–4.

[20] Seeman N. Nucleic acid junctions and lattices. J. Theoretical Biology 1982;99:237–47.

[21] Han D et al. Single-stranded DNA and RNA origami. Science 2017;385:1402.

[22] He Y et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. Nature 2008;452:198–201.

[23] Chenxiang L et al. DNA tile based self-assembly: building complex nanoarchitectures. ChemPhysChem 2006;7:1641–7.

[24] Zheng J et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. Nature 2009;461:74–7.

[25] Ke Y et al. Three-dimensional structures self-assembled from DNA bricks. Science 2012;338:1177–83.

[26] Ong LL et al. Programmable self-assembly of three-dimensional nanostructures from 10^4 unique components. Nature 2017;552:72–7.

[27] Rothemund PWK. Folding DNA to create nanoscale shapes and patterns. Nature 2006;440:297–302.

[28] Chanseok L et al. Polymorphic design of DNA origami structures through mechanical control of modular component. Nat Commun 2017;8:2067.

[29] Suping L et al. A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. Nat Biotech 2016;36:258–64.

[30] Ora A et al. Cellular delivery of enzyme-loaded DNA origami. Chem Comm 2016;52:14161.

[31] Younghyun C, Jong BL, Jinkee H. Controlled release of an anti-cancer drug from DNA structured nano-films. Sci Rep 2014;4:4078.

[32] Angelin A et al. Protein-functionalized DNA nanostructures as tools to control transcription in zebrafish embryos. Chem Open 2016;6:33–9.

[33] Zhang Z et al. Placing and shaping liposomes with reconfigurable DNA nanocages. Nat Chem 2017;9:653–9.

[34] Ramakrishnan S et al. Cation-induced stabilization and denaturation of DNA origami nanostructures in urea and guanidinium chloride. Small 2017;13:1702100.

[35] Martin TG et al. Design of a molecular support for cryo-EM structure determination. PNAS 2016;113:E7456–63.

[36] Wagenbauer KF, Sigl C, Dietz H. Gigadalton-scale shape-programmable DNA assemblies. Nature 2017;552:78–83.

[37] Agarwal NP et al. Structural transformation of wireframe DNA origami via DNA polymerase assisted gap-filling. ACS Nano 2018;12:2546–53.

[38] Zhang F et al. Complex wireframe DNA origami nanostructures with multi-arm junction vertices. Nat Nanotech 2015;10:779–84.

[39] Matthies M et al. Triangulated wireframe structures assembled using single-stranded DNA Tiles. ACS Nano 2019;13:1839–48.

[40] Slone SM et al. Molecular mechanics of DNA bricks: in situ structure, mechanical properties and ionic conductivity. New J Phys 2016;18:055012.

[41] Pettersen EF et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004;13:1605–12.

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

[43] Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph 1996;14(1):33–8.

[44] Williams S et al. Tiamat: a three-dimensional editing tool for complex DNA structures. DNA Comput 2008;1:90–101.

[45] Jun H. et al. Autonomously designed free-form 2D DNA origami, Sci Adv 2019:5;eaav0655.

[46] De LIano E. et al. Adenita: interactive 3D modeling and visualization of DNA, Nanostructures bioRxiv 2019: 10.1101/849976,

[47] Ouldridge TE, Louise AA, Doye JPK. DNA nanotweezers studied with a coarse-grained model of DNA, Phys Rev Lett 2010:104;178101.

[48] Ouldridge TE, Louise AA, Doye JPK. Structural, mechanical and thermodynamic properties of a coarse-grained DNA model, J Phys Condens Matter 2012:22;104102.

[49] Phillips JC et al. Scalable molecular dynamics with NAMD. J Comput Chem 2005;26:1781–802.

[50] Pan K et al. Lattice-free prediction of three-dimensional structure of programmed DNA assemblies. Nat Commun 2014;5:520–4.

[51] Doye JPK et al. Coarse-graining DNA for simulations of DNA nanotechnology, Phys Chem Chem Phys 2013:15;20395.

[52] Maffeo C, Yoo J., Aksimentiev A. De novo reconstruction of DNA origami structures through atomistic molecular dynamics simulation, Nucleic Acids Res 2016;44:3013–3019.

[53] Nakano S et al. Nucleic acid duplex stability: influence of base composition on cation effects. Nucleic Acids Res 1999;14:2957–65.

[54] Owczarzy R et al. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. Nucleic Acids Res 2008;36:W163–9.

[55] Martin TG, Dietz H. Magnesium-free self-assembly of multi-layer DNA objects. Nat Commun 2012;3:1103.

[56] Wagenbauer KF et al. How we make DNA origami. ChemBioChem 2017;18:1873–85.

[57] Zhang Z et al. Self-assembly of DNA origami and single-stranded tile structures at room temperature. AngewChem 2013;52:9213–23.

[58] Ahmadi Y et al. The Brownian and flow-driven rotational dynamics of a multicomponent DNA origami-based rotor. Small 2020:e2001855.

[59] Song J et al. Reconfiguration of DNA molecular arrays driven by information relay. Science 2017;357:6349.