Supplementary material: Direct observation and rational design of nucleation behavior in addressable self-assembly

Martin Sajfutdinow,[1,2] William M. Jacobs,[3] Aleks Reinhardt,[4] Christoph Schneider,[1] and David M. Smith[1]
[1)] *Fraunhofer Institute for Cell Therapy and Immunology IZI, Department of Diagnostics, Perlickstraße 1, 04103 Leipzig, Germany*
[2)] *Faculty of Chemistry and Mineralogy, Leipzig University, Johannisallee 29, 04103 Leipzig, Germany*
[3)] *Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, United States*
[4)] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom*

(Dated: 9 April 2018)

1 EXTENDED METHODS

If not otherwise noted, chemicals were purchased from Sigma Aldrich Chemie GmbH (Munich, Germany).

1.1 Annealing protocols

DNA sequences for the target cuboid structure are given in Sec. SI-5. For the all-BB system, sequences were taken from Ref. 1 (Table S8), and DNA strands were appropriately decoupled to split the relevant boundary bricks for the face-BB, edge-BB and no-BB systems. All sequences were purchased from Eurofins Genomics in 100 µM stocks in ddH_2O, and then pooled using a Tecan Genesis Workstation 150 liquid handling robot. We used a strand concentration of 153 nM in 1×assembly buffer, i.e., a solution of 15 mM $MgCl_2$, 0.5 mM EDTA and 5 mM Tris, pH 8. The strand solution was denatured at 90 °C for 10 min and then gradually cooled. We used two linear cooling protocols: (i) in the 15.2 h protocol, the reciprocal cooling rate was 12 min K^{-1}, and (ii) in the 66 h protocol, it was 52 min K^{-1}. The annealed samples were stored at 4 °C.

Prior to the reference DLS measurement (positive control), the all-BB sample assembled in the 66 h protocol was supplemented with 2.5 mM EDTA to reduce high-molecular-weight contamination. In the context of DLS experiments, we therefore refer to this sample as 'purified'.

Prior to the reference AFM imaging (positive control), the all-BB sample assembled in the 66 h protocol was ultrafiltered using Amicon (UFC510024, Millipore, Merck KGaA, Darmstadt, Germany) filter units, with a molecular weight cut-off of 100 kDa, to reduce the fraction of small particle contamination and to improve the image quality. To this end, the assembled sample was mixed with pre-chilled 1 × assembly buffer to the maximum admitted volume and centrifuged for 10 min at 14 000*g* at 4 °C. Subsequently, the flowthrough was discarded, and the filter unit was loaded with buffer and centrifuged again. This process was repeated three times in total. Finally, the concentrated filtrate was eluted at 1000*g* for 2 min at 4 °C.

1.2 Atomic force microscopy

Samples were prepared following the 66 h annealing protocol. A freshly cleaved mica disc was coated with 100 µL of 0.5 wt% poly-L-ornithine solution for 5 min and rinsed three times with 1 × assembly buffer. In order to be able to image samples in liquid mode, an acrylic glass ring was glued by Thin Pour (Reprorubber) onto a slide to surround the mica disc and form a fluid cell. For each sample, 1.5 pmol per brick was deposited on the coated mica for 10 min. Afterwards, the cell was filled with 1×assembly buffer and imaged using the JPK Nanowizard 3 atomic force microscope and a BioLever Mini cantilever in intermittent contact mode in liquid. Images were recorded with a target amplitude of 15 nm.

Quenching experiments, designed to stop the hybridization reaction at a given temperature during the annealing protocol, were done by immobilizing 5 µL of the reaction mixture on poly-L-ornithine coated and pre-equilibrated mica discs. As a negative control, the same procedure was performed using a random selection of ssDNA strands that did not contain complementary sequences.[2] For both the all-BB structure and the negative control, samples were quenched from 318 K, 314 K, 311 K, 308 K and 300 K during annealing protocol (i) and imaged by AFM, as described above, at ambient temperature.

1.3 Agarose gel electrophoresis

Assembly of DNA brick structures was confirmed by non-denaturing agarose gel electrophoresis. Samples (300 fmol per brick) were analyzed on a gel made from 2 wt% agarose in 0.5 × TBE and 10 mM $MgCl_2$. Electrophoresis was performed at 80 V and 4 °C for 2 h. The gel was post-stained with 0.5 µg mL^{-1} ethidium bromide solution and scanned in using the Intas GDS gel set instrument for structure visualization. To estimate structure yield, the band intensity was approximated by fitting densitometry profiles with an SQP algorithm to Gaussian functions using the GelBandFitter software.[3] The mass of the structure fractions was estimated via a 1 kb standard (GeneRuler, ThermoFischer Scientific) and related to the total mass loaded of 850 ng.

1.4 Static and dynamic light scattering

The same conditions as in the 15.2 h annealing protocol were used and the measurement was performed in the last 2 min of the 12 min cooling step. 20 µL samples were filled into ZEN2112 quartz cuvettes (Malvern), covered by molecular biology grade mineral oil, and sealed with a plastic lid that was further fixed with tape. Light scattering was measured using a Malvern Zetasizer NanoZSP apparatus at an angle of 173°. The viscosity of the samples was determined at five temperatures spanning the region of interest and fitted to $\eta/(10^{-5}\,\mathrm{Pa\,s}) = 1.78\times\exp[617/(T/\mathrm{K}-138.5)]$. The refractive index was measured to be 1.331.

For dynamic light scattering, the intensity auto-correlation

function was computed from 12 measurements at 10 s intervals. We interpreted the DLS data in the dilute limit by assuming that all particles diffused independently of one another, since the total strand concentration (approximately 40 μM) implies that single strands ($R_h \sim 2.7$ nm) in solution occupied a volume fraction of approximately 0.2 %. We further assumed that, after the initial equilibration period of 10 min, the distribution of cluster sizes remained nearly constant over the DLS measurement period at the end of each temperature step; this assumption is consistent with our observations of rate-limiting structure nucleation.

When analyzing DLS data for solutions comprising a range of particle sizes, the inverse Laplace transform used to obtain a particle size distribution from the intensity auto-correlation function is not uniquely determined,[4] and the choice of fitting functions and parameters can affect the final result.[5] We have therefore computed multiple fits to the distribution of hydrodynamic radii using several regularization methods, including a range of different smoothing exponents and a maximum entropy constraint (Fig. S3). Although the agreement is not perfect for the individual data points, the trends for the distributions and the fits to a linear combination of Gaussian functions are nearly the same regardless of the regularization procedure used, indicating that the conclusions drawn from the DLS data are robust with respect to the choice of regularization method. We chose to use the smoothness constraint regularization method recommended in Ref. 5 with a smoothing exponent of $m = 8/5$ (corresponding to Fig. S3b) for all data reported in the main text.

1.5 Fluorescence annealing

In the fluorescence annealing experiments, the same conditions as in the 15.2 h annealing protocol were used, except that 10 nM SYBR green I solution[6] was added to the strand mixture. SYBR green I in buffer solution was analyzed as a negative control. Samples were placed on a MicroAmp Fast Plate 96-well tray and sealed with adhesive film. The plate was loaded onto the ABI Prism 7900HT-Fast Real Time PCR system, with dye excitation effected by an argon ion laser at 488 nm. The fluorescence signal was detected at 525 nm every 8.5 s and averaged over time at each temperature, and its derivative with respect to temperature was computed numerically. The data were smoothed via a Gaussian filter with a standard deviation of 1.5 K.

2 MONITORING STRAND HYBRIDIZATION

2.1 Fluorescence measurements

We monitored the progress of domain hybridization during the annealing protocol via fluorescence, using SYBR green I as a double-stranded DNA probe (see Sec. 1.5).[7] We observed a dominant maximum in the fluorescence derivative between 335 K and 350 K for all structures with boundary bricks (Fig. S4a), indicating a significant amount of base-pairing at relatively high temperatures. However, as we discuss in the main text, no complete structures were assembled at these temperatures.

Comparison with theoretical annealing curves suggests that the assembly of boundary-brick structures is a two-step process. To demonstrate this, we show in Fig. S4b the temperature derivative of the equilibrium number of base pairs in a solution of monomers and dimers,[8,9] assuming that stable misbonding between non-complementary domains cannot occur (see below). The high-temperature transitions correspond to the hybridization between pairs of boundary bricks (where continuous 24-bp segments are hybridized) or between one scaffold strand and one boundary brick (with 16-bp hybridized segments). Consequently, the assembly of the full structure must occur in the presence of these pre-formed clusters. These calculations also indicate that the fluorescence-signal contributions from each domain length overlap significantly, since the domain melting temperatures vary widely according to their specific sequences, and each hybridization reaction tends to occur over a broad ($\gtrsim$10 K) range of temperatures. In particular, the theoretical annealing curves predict a broad maximum associated with the 8-bp domains near 295 K.

Analysis of fluorescence data has previously been used to distinguish between single- and multi-step assembly mechanisms for DNA tile systems with varying domain lengths. For example, a similar step-wise assembly process was seen in DX-tile structures comprising short (10- and 11-bp) and long (21-bp) hybridizations, and the presence of two distinct maxima in the fluorescence derivative was interpreted as evidence of hierarchical assembly.[10] By contrast, fluorescence measurements of DNA-brick crystallization using equal-length domains exhibited no evidence of hierarchical self-assembly.[11] In our measurements, there appear to be multiple local maxima in the annealing curves at temperatures below the scaffold-strand $T_m \simeq 315$ K, the highest temperature at which our theoretical calculations predict that a lattice of scaffold strands can be thermodynamically stable. However, these signals are significantly weaker than the higher-temperature hybridizations which dominate the fluorescence signal. Interpreting the lower-temperature maxima is additionally hindered by several known sources of bias, including high background signals[6] and the preferential binding of SYBR green I to GC-rich sequences.[12] Furthermore, the intercalating SYBR green I probes distort the double-helical structure of DNA molecules,[13] which increases their melting temperatures[14] and precludes a quantitative analysis.

2.2 Hybridization calculations

All hybridization calculations were carried out using the SantaLucia parameterization and the solution conditions described in the Methods section of the main text. In this section, we consider a two-state model (i.e. bonded or not bonded) for each domain and examine the simple case where pairs of strands hybridize to form dimers, but not larger multimers. We denote the hybridization free energy between complementary domains on a pair of strands i and j by ΔG_{ij}. The equilibrium probability that a strand i is correctly hybridized with its putative neighbor strand j is

$$p_{ij}(T) = \frac{\rho \exp\left(-\Delta G'_{ij}/k_B T\right)}{1 + \rho \exp\left(-\Delta G'_{ij}/k_B T\right)}, \qquad (1)$$

where ρ is the dimensionless strand number density, k_B is the Boltzmann constant, T is the absolute temperature, and we assume that all species are present in equal concentrations. The hybridization free energies are written as $\Delta G'_{ij}$ to indicate that we use the longest complementary subsequence of strands i and j, which, due to the random sequence design, is occasionally longer than the intended domain length. To calculate the total change in base-pairing during an annealing protocol (Fig. S4b), we took the temperature derivative of the ensemble average of

correctly formed base pairs,

$$\text{Hybridization} = -\frac{\mathrm{d}}{\mathrm{d}T}\sum_{\substack{i<j\\ j\in\mathcal{E}(i)}} l_{ij}p_{ij}(T), \tag{2}$$

where l_{ij} is the length of each hybridizing domain.

3 CLUSTER POPULATION RATIOS

We assume that annealing is slow, so that nucleation is always rate-limiting. We can write the nucleation barrier height as

$$F^\dagger = -n^\dagger k_\mathrm{B}T\ln\rho_\mathrm{m} + \varepsilon(T)E^\dagger + C, \tag{3}$$

where $n^\dagger$ is the number of independent subunits in the critical nucleus, $E^\dagger$ is the number of 8-bp bonds in the critical nucleus, and C is a constant that accounts for the (effective) number of parallel nucleation pathways, as well as the rotational entropy terms. The bond energy ε is a decreasing function of temperature, while the per-species monomer concentration ρ_m, indicating the number of monomers per unit volume, also decreases as the reaction progresses. Initially, we have ρ_T of each species. For simplicity, let us assume that, given this initial monomer concentration, the barrier is infinitely high above some critical temperature T_0. Nucleation begins once $T \le T_0$, where $F^\dagger$ is finite. (In reality, nucleation can begin as soon as the target structure, or any large cluster, becomes thermodynamically stable. However, the nucleation rate is proportional to $\exp(-F^\dagger/k_\mathrm{B}T)$, so the highest barrier that can be crossed depends on the cooling rate.)

Nucleation will proceed at a given temperature until ρ_m decreases to a point where $F^\dagger$ is again insurmountable. Denoting this critical barrier height by $F_0^\dagger$, we can relate the final monomer concentration ρ_m at any temperature to the initial concentration at the critical temperature,

$$F_0^\dagger - C = -n^\dagger k_\mathrm{B}T\ln\rho_\mathrm{m}(T) + \varepsilon(T)E^\dagger \tag{4}$$

$$= -n^\dagger k_\mathrm{B}T_0\ln\rho_\mathrm{T} + \varepsilon(T_0)E^\dagger, \tag{5}$$

so that

$$\frac{\rho_\mathrm{T}}{\rho_\mathrm{m}} = \exp\left\{-\frac{E^\dagger}{n^\dagger}\left[\frac{\varepsilon(T)}{k_\mathrm{B}T} - \frac{\varepsilon(T_0)}{k_\mathrm{B}T_0}\right]\right\}. \tag{6}$$

Assuming that the intensity of each peak is proportional to the concentration of unassembled strands (m) or assembled structures (c), respectively, the ratio of the scattering intensities is

$$\frac{I_\mathrm{c}}{I_\mathrm{m}} = \frac{R_\mathrm{h,c}^6}{NR_\mathrm{h,m}^6}\left(\frac{\rho_\mathrm{T}-\rho_\mathrm{m}}{\rho_\mathrm{m}}\right) \tag{7}$$

$$= \frac{R_\mathrm{h,c}^6}{NR_\mathrm{h,m}^6}\left(\exp\left\{-\frac{E^\dagger}{n^\dagger}\left[\frac{\varepsilon(T)}{k_\mathrm{B}T} - \frac{\varepsilon(T_0)}{k_\mathrm{B}T_0}\right]\right\} - 1\right), \tag{8}$$

where N is the number of distinct subunits in the target structure. Furthermore, because $\varepsilon/k_\mathrm{B}T$ is a nearly linear function of T in the range of interest (Fig. S9), we expect the intensity ratio to have a functional form

$$\frac{I_\mathrm{c}}{I_\mathrm{m}} = \text{const}\times\left\{\exp[-a(T-T_0)] - 1\right\}, \tag{9}$$

where $a = (E^\dagger/n^\dagger)(\mathrm{d}\beta\varepsilon/\mathrm{d}T)$ and $\beta = 1/k_\mathrm{B}T$. Using a linear fit to the energies as a function of temperature at temperatures of interest (Fig. S9), $\mathrm{d}\beta\varepsilon/\mathrm{d}T \approx 0.34\,\mathrm{K}^{-1}$. From the theoretical free-energy profiles shown in the main text, we know that for edge BBs, $E^\dagger/n^\dagger = 7/6$, whilst for face BBs, the ratio is $6/5$. Hence we can estimate that $1/a \approx 2.5\,\mathrm{K}$.

To calculate the intensity associated with each peak in the DLS data, we first fitted a sum of Gaussians to the distribution function, $f(R_\mathrm{h})$. We then numerically integrated the peak associated with the Gaussian function f_g, according to

$$I_\mathrm{c/m} = \int_0^\infty \min\left[f_\mathrm{g,c/m}(R_\mathrm{h}),\ f(R_\mathrm{h})\right]\mathrm{d}\ln R_\mathrm{h}. \tag{10}$$

In reality, the appearance of aggregates at low temperatures, which tend to increase the mean R_h of the assembled population, means that the ratio of the scattering intensities is not exactly proportional to the ratio of the cluster concentrations. However, this effect is relatively small over the range of temperatures of interest (approximately 8 kelvin below T_0; see Fig. 4a). Instead, the exponential increase in the intensity ratio as a function of decreasing temperature shown in Fig. 5c is driven primarily by an exponential decrease in the scattering intensity of the unassembled population upon cooling below T_0. Such behavior is consistent with the theoretically predicted evolution of the unassembled-strand population shown in Fig. 5b.

4 SUPPLEMENTAL FIGURES

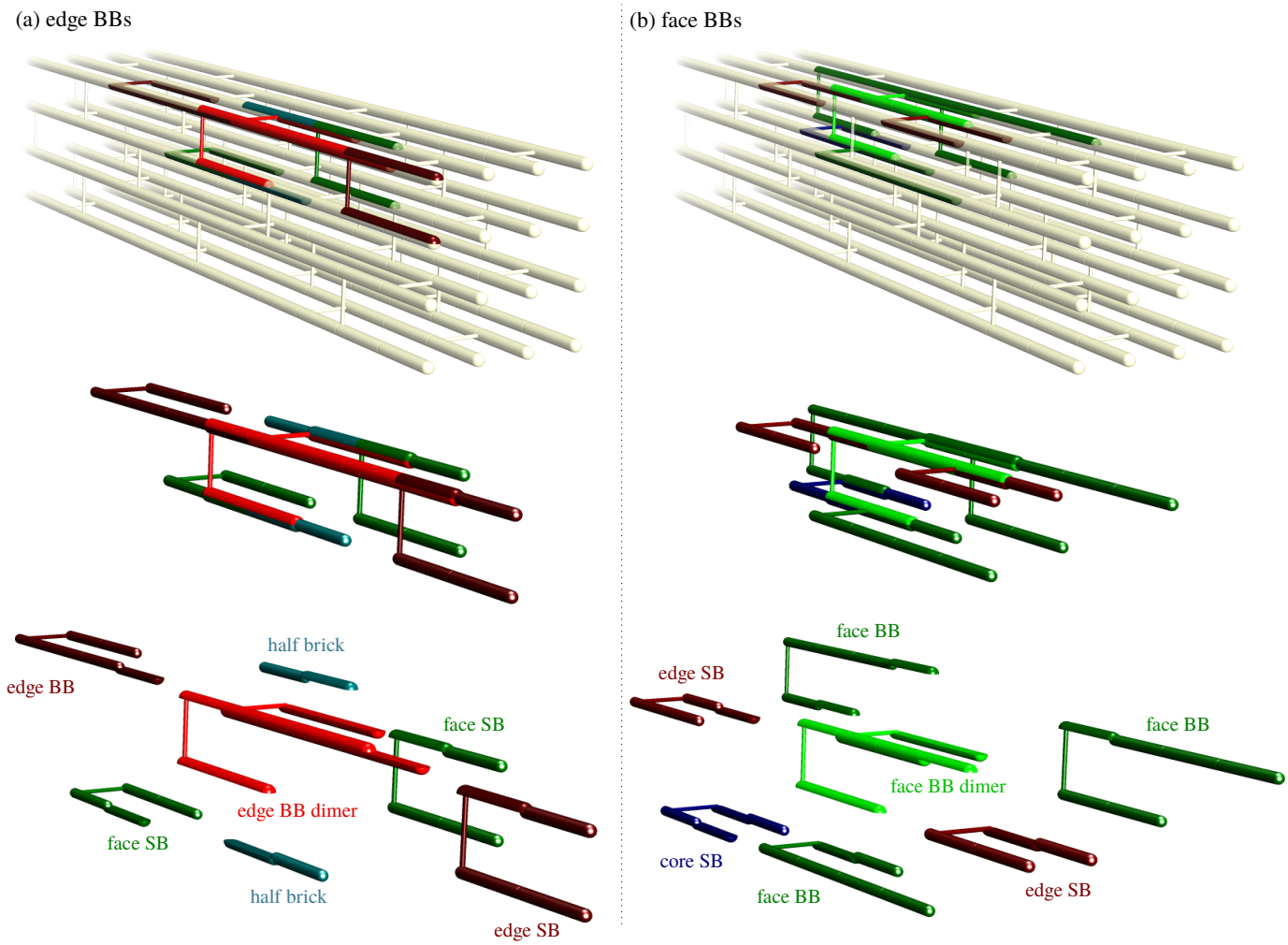

FIG. S1. Boundary brick dimers and their nearest neighbors in a schematic representation. In the top panel, the location within the target structure is shown. In the bottom panel, the neighbors are shown spread out and labeled to make their identification clearer. 'SB' stands for a 32-nt 'scaffold brick'. The edge-BB system's nucleation properties were also investigated by merging some bricks, as described in the main text. In particular, the 'merged-A' building block corresponds to the edge BB dimer shown in red. The 'merged-B' building block corresponds to the edge BB dimer and one of the face SBs shown in dark green. Either one of these face SBs could have been chosen, as both of them have direct connections to core strands. In our simulations, the face SB that is merged with the edge BB dimer is the one whose center of mass is nearer the cuboid's principal axis in the target structure.

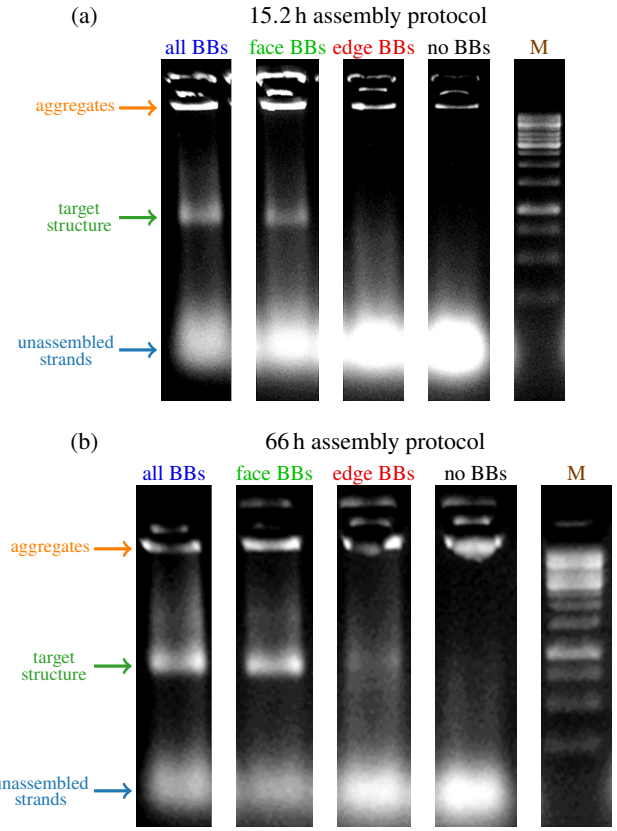

FIG. S2. Gel electrophoresis of samples at the end of (a) 15.2 hour or (b) 66 hour linear assembly protocols. Samples were assayed in 2 % agarose gel. Lane M contains a GeneRuler 1 kb ladder that was used to reference the assembly yield. The bands corresponding to the target structures, unassembled strands and aggregates are indicated.

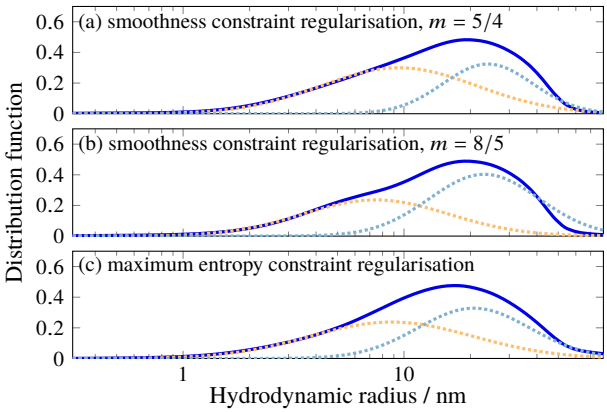

FIG. S3. Example size distribution functions (solid blue lines) for the all-BB system at 310 K determined using three regularization methods. We used a smoothness constraint functional[4] with smoothing exponents[5] of (a) $m = 5/4$ and (b) $m = 8/5$, as well as (c) a maximum entropy constraint with a Gaussian prior distribution.[5] Dotted lines show the Gaussian functions determined by fitting a linear combination of Gaussians to each distribution function.

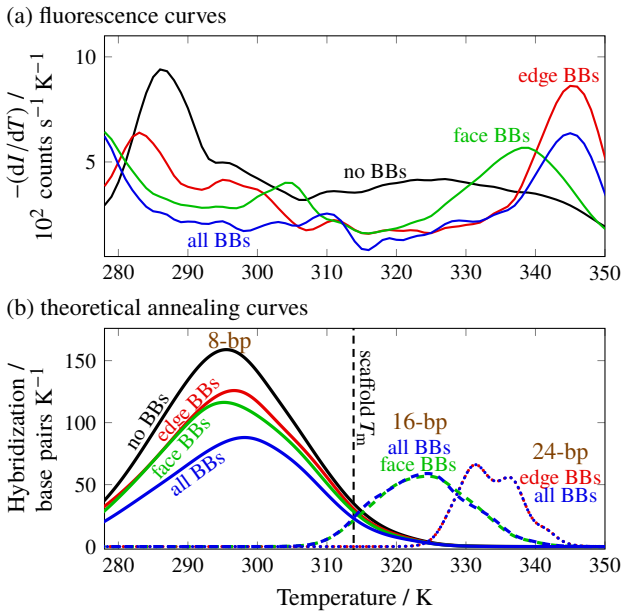

FIG. S4. (a) The derivative of the fluorescence signal I with respect to the temperature obtained from a 15.2 h annealing protocol. (b) The corresponding theoretical annealing curves (see Sec. 2.2). In agreement with the experimental fluorescence data, contributions from 16- (dashed lines) and 24-bp (dotted lines) hybridizations dominate at higher temperatures. The predicted scaffold-strand T_m is also shown.

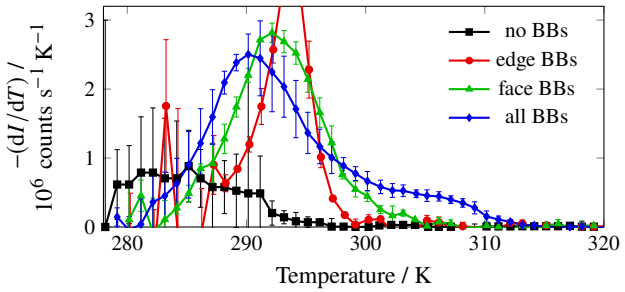

FIG. S5. The derivative of the static light scattering intensity I with respect to temperature for self-assembly following the 15.2 h annealing protocol.

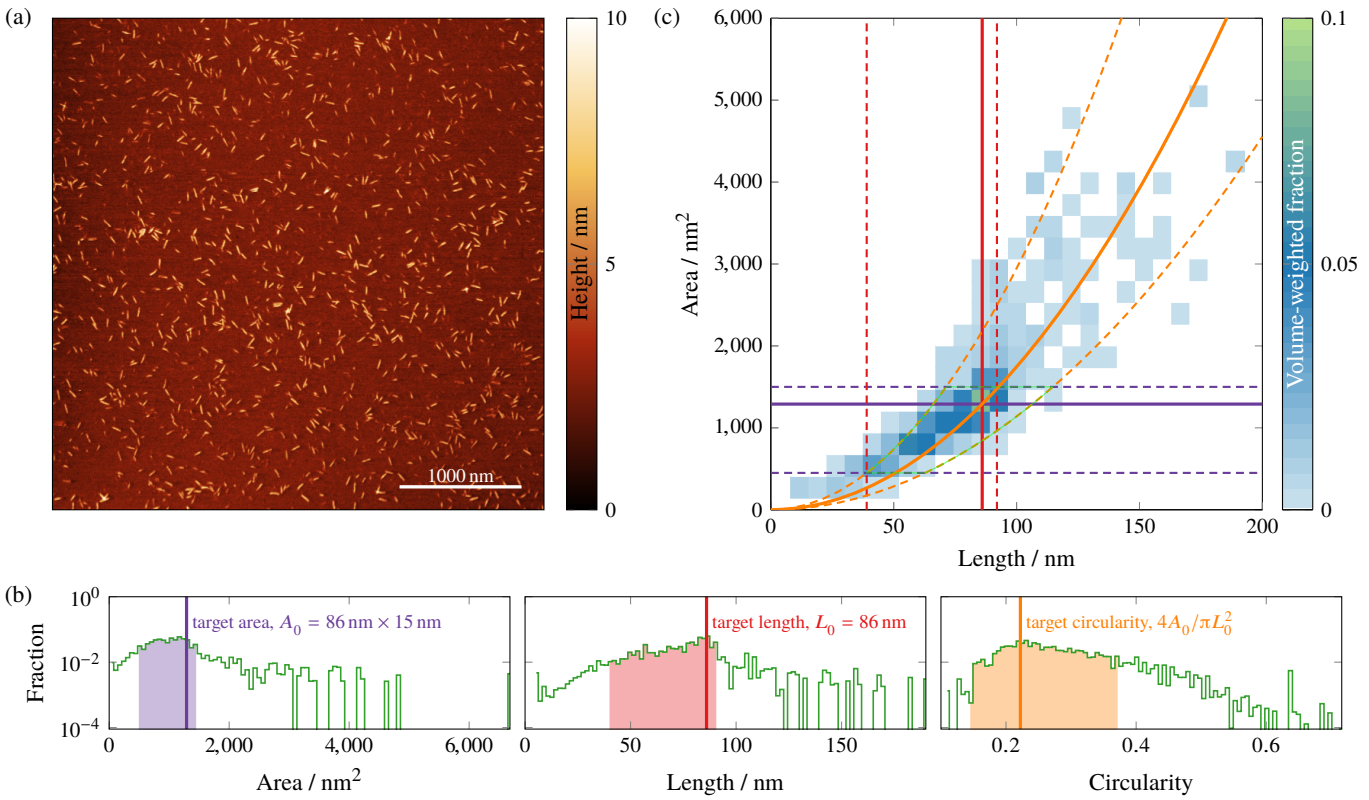

FIG. S6. Criteria for identifying target structures via AFM. (a) AFM image of the positive-control all-BB structures (see Sec. 1.1). (b) Area, length and circularity distributions of the particles identified in this image (see Methods). All distributions are weighted by the particle volume to prevent tiny particles from skewing the distributions. The solid vertical lines show the expected values for an ideal target structure when treating the cuboid as a cylinder with diameter $d \sim 15$ nm. (c) Two-dimensional volume-weighted distribution of imaged particle areas and lengths. The solid lines indicate the expected values for an ideal target structure, while the dashed lines correspond to the boundaries of the shaded regions in panel (b); the parabolic curves show the circularity, $4A/\pi L^2$. Notably, the peak of this distribution (green square) coincides with the expected area, length and circularity. Based on this distribution, we chose to use the area and circularity criteria ($450\,\text{nm}^2 \leq A \leq 1500\,\text{nm}^2$ and $0.145 \leq 4A/\pi L^2 \leq 0.375$, indicated by the translucent green lines) to identify particles as correctly assembled structures. Using these criteria, the AFM-determined yield of the positive control is 53 %.

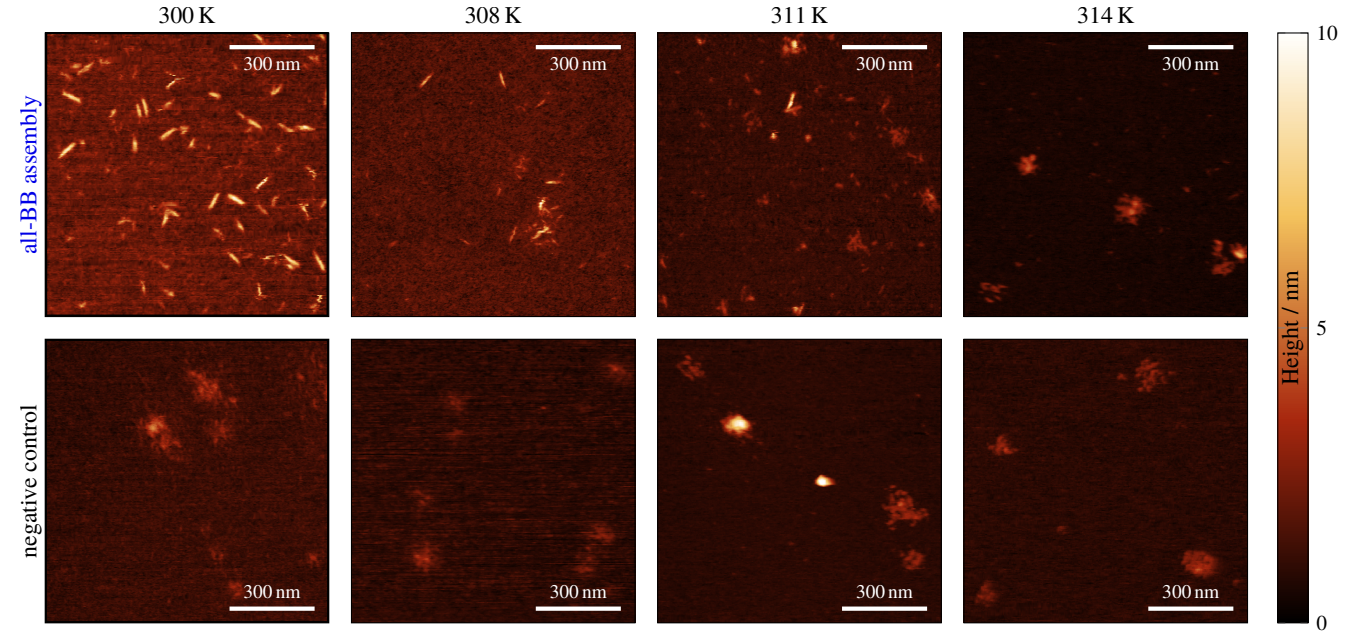

FIG. S7. AFM images for all-BB structures (top row) and negative controls (bottom row). Samples were prepared by rapidly quenching the system from the indicated temperature (column labels) during the annealing protocol (see Sec. 1.2). The negative control consists of a collection of similar-length oligonucleotides that were not designed to have complementary sequences. The bright amorphous clusters seen in the negative control and the all-BB structures quenched from high temperatures indicate large aggregates of oligonucleotides that form during quenching.

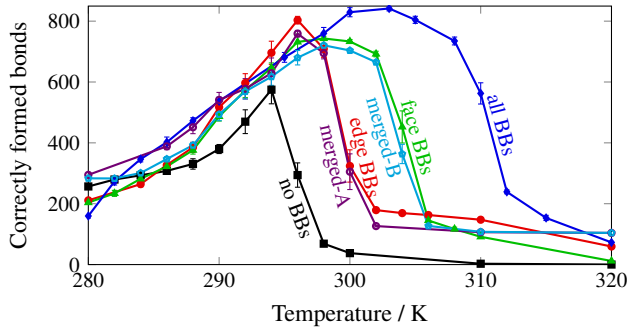

FIG. S8. The number of correctly formed bonds in the system as a function of temperature from Monte Carlo simulations. Each data point corresponds to an average over ten independent simulations in the long time limit once nucleation has occurred. Error bars give the standard deviation in each case.

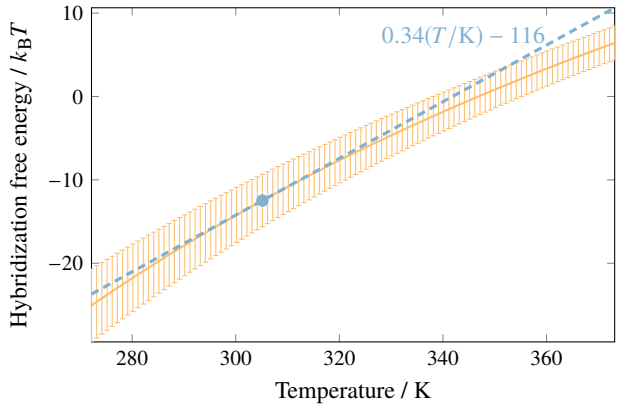

FIG. S9. The mean hybridization free energies of 8-bp interactions as a function of temperature for the no-BB structure, computed via the SantaLucia thermodynamic model,[8] with error bars reflecting the standard deviation. The tangent to the curve at 305 K is also shown, demonstrating that the hybridization free energy is well described by a linear function over the region of interest (295 K to 315 K).

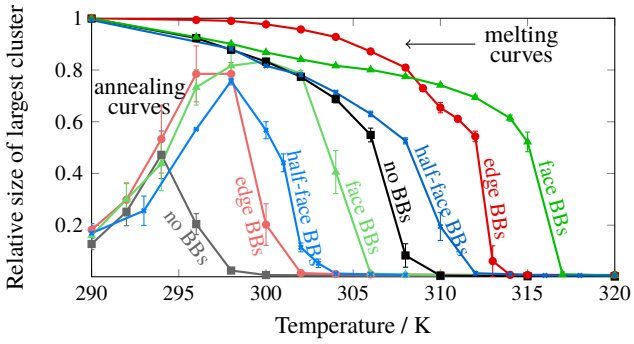

FIG. S10. Melting and annealing curves for the half-face BB structure in Monte Carlo simulations (cf. Fig. 4c and Fig. 6b).

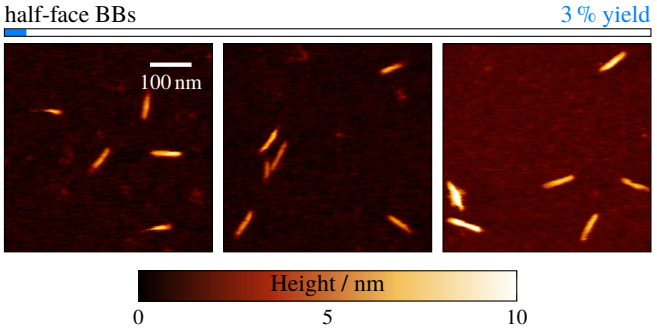

FIG. S11. AFM images and the yield, as determined by gel electrophoresis, for the half-face BB structure.

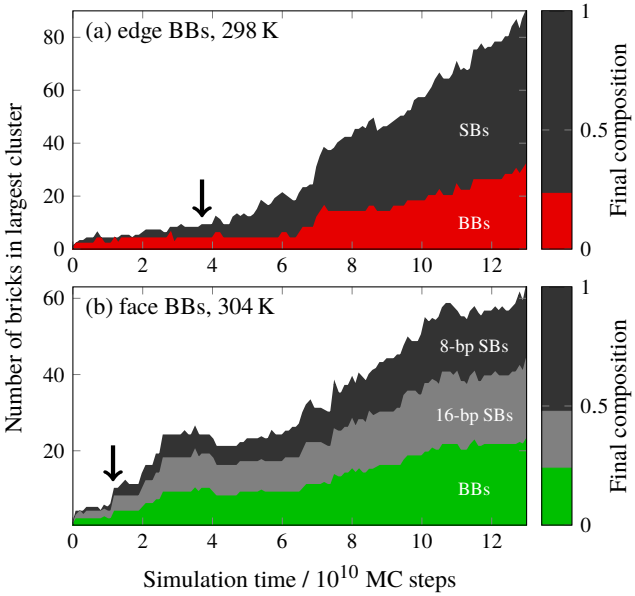

FIG. S12. The size of the largest cluster as a function of time in representative Monte Carlo trajectories for the (a) edge-BB and (b) face-BB systems. The composition of the largest cluster is shown in terms of boundary bricks (BBs) and scaffold bricks (SBs). In the face-BB system, the scaffold bricks are further divided into those that have only 8-bp hybridizations with neighboring bricks and those which form 16-bp hybridizations with face BBs. The relative proportions of each type of brick in the target structure are shown in the right-hand panel in each case. Both trajectories start at roughly the point where a fluctuation leads to nucleation and subsequent growth. Arrows indicate approximately where the theoretically predicted post-critical nucleus (see Fig. 7) first appears in each trajectory. In (a), this cluster has two BB dimers (four BBs) and four SBs. In (b), this cluster has three BB-SB dimers (three BBs and three 16-bp SBs) and two 8-bp SBs. In both cases, the cluster size increases rapidly after this nucleation event. BBs dominate in the early part of the growth phase shown here, where they comprise a greater proportion of the post-critical clusters than they do in the final assembled structure.

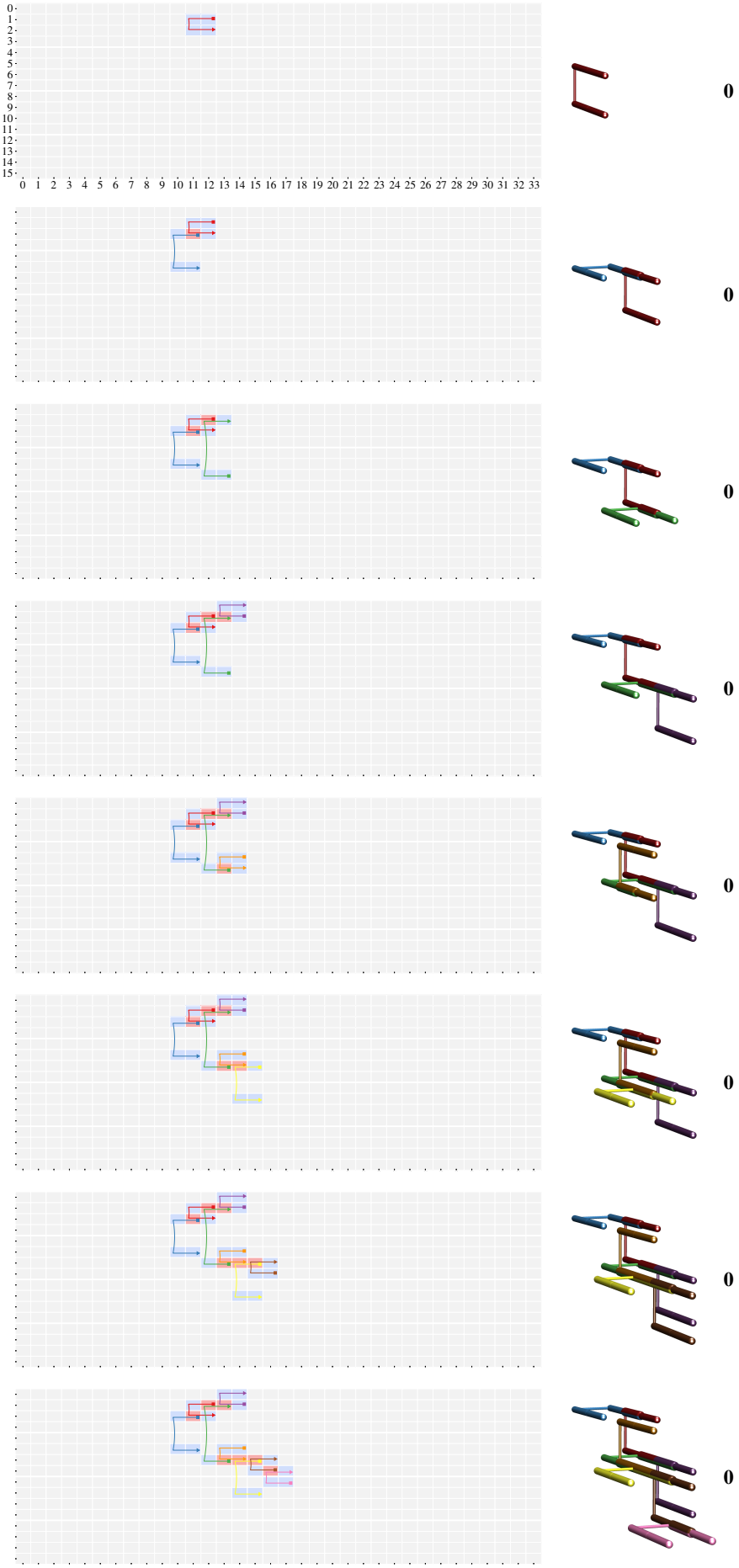

FIG. S13. Example low free-energy nucleation pathway for the no-BB system in two representations: a Cadnano-style connectivity diagram and a three-dimensional rendering. In the latter, DNA brick domains are represented by cylinders. Non-bonded domains are represented by smaller cylinders, while where two DNA bricks are bonded, larger multicolored cylinders are used. Each new monomer or multimer added to the cluster along the nucleation pathway is colored in a different hue. The bold number to the right of each structure indicates the number of multimers in the structure.

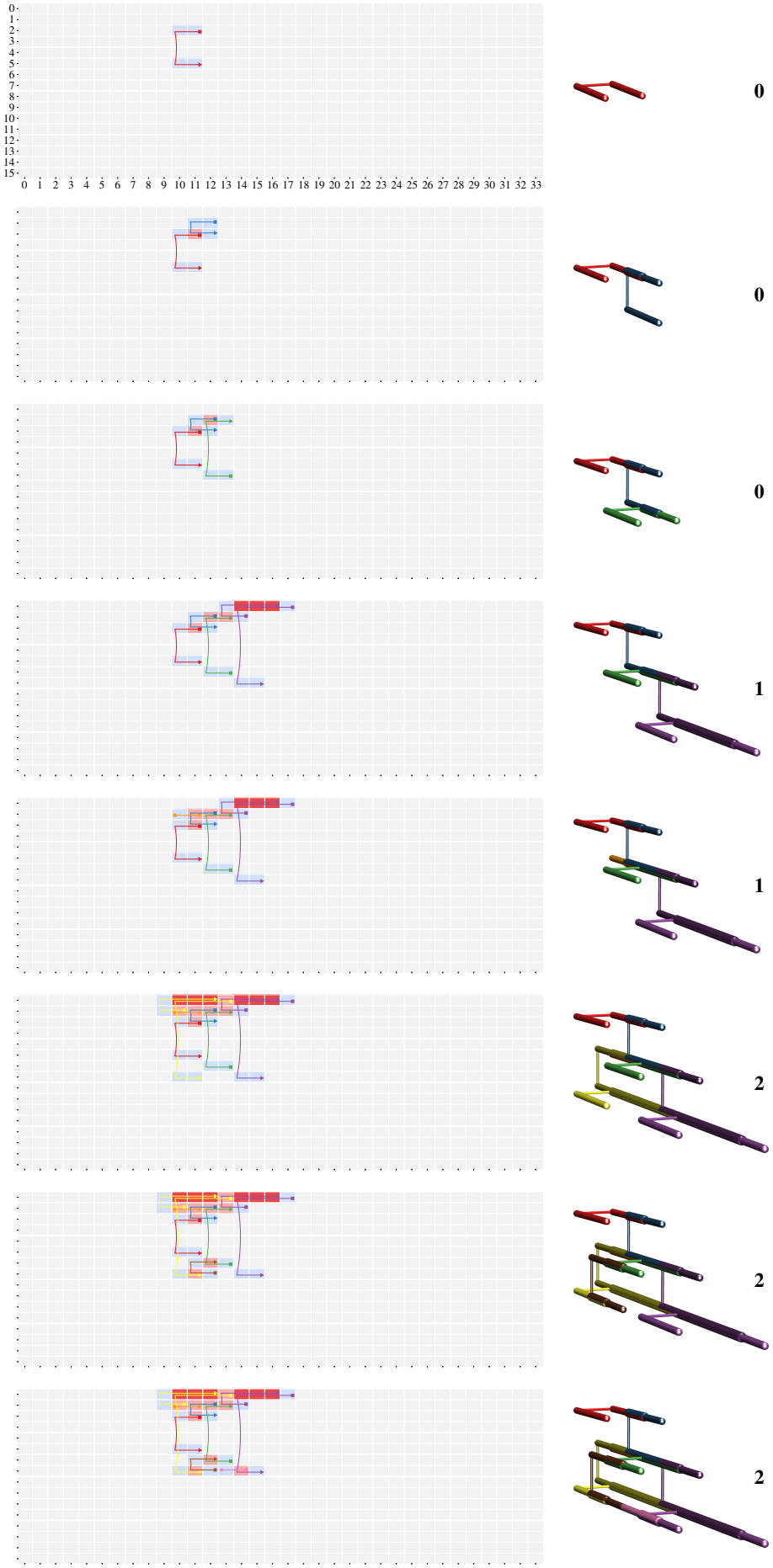

FIG. S14. Example low free-energy nucleation pathway for the edge-BB system in two representations, as in Fig. S13.

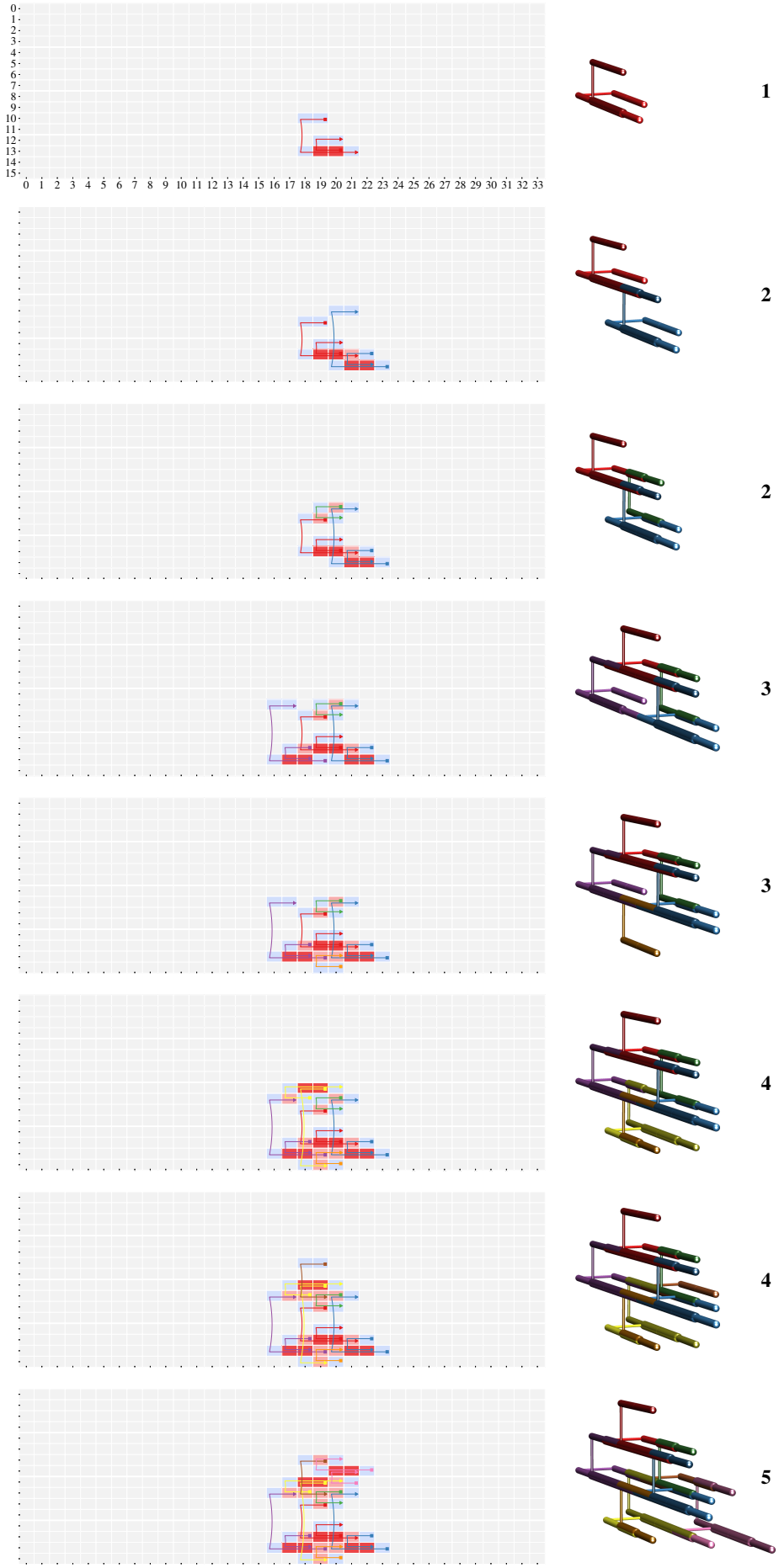

FIG. S15. Example low free-energy nucleation pathway for the face-BB system in two representations, as in Fig. S13.

5 DNA BRICK SEQUENCES

The following sequences comprise our library of DNA bricks.

ID	Sequence
1	TTCTTAAA TTTTTTTT
2	TTGGCTGA TTTTTTTT
3	TTTTTTTT CCGCGTAA
4	TTTTTTTT CCAACAGG
5	TATGGTGA TTTTTTTT TTTTTTTT TGGCACCC
6	ATCCGAAG TTTTTTTT TTTTTTTT CGCGGACA
7	CGCTGATC TTTTTTTT TTTTTTTT GGGGATGA ATCCTTCC AACATCTC
8	TCTGATAT TTTTTTTT TTTTTTTT GCACTGCC CCTACTCT CACCTTTC
9	ATATCAGA GAGCACAG
10	TCCGGGAT GGGTGCCA
11	AGAGTCTG GAGATGAT
12	GCTTCGCG CCTGTTGG
13	TAATATAA TGTCCGCG GATCAGCG AATTCGAC
14	AGTTAGCA TTACGCGG GTACCCTG GACCGCTA
15	GGAAGGAT TCATCCCC TTTAAGAA TGATCGCA CCAGGTTA AGTGGCTC
16	TACATCTC GCTCGTCC TCACCATA ATCCACCA CCCATGCA GAAAGATA
17	GTGCGCTT CTGTGCTC ATCCCTTG CTCCCGAA
18	TCCCTAAA GTCGAATT GAGATGTA AGTTTCAC
19	ACCTTTTA ATCATCTC CAGACTCT TTTTTTTT TTTTTTTT GAGCCGAT
20	GGGAGATT TAGCGGTC CAGGGTAC TTTTTTTT TTTTTTTT GGACGAGC
21	CAGTCTTT GAGATGTT GAACATCT TTTGGTTG
22	AAGGTAAT GTGAAACT CAATGATA TGACGGAT
23	ACATCTCG GAAAGGTG AGAGTAGG GGCAGTGC TCAGCCAA ACAGTGGG
24	CTTACACT TTCGGGAG CAAGGGAT ATCGGCTC CTTCGGAT TCTCTGGC
25	TATCATTG GCCAGAGA TGCTAACT GTATGACG GACCAGGT CTCCTAGG
26	TGCATGGG TGGTGGAT CGCGAAGC TGTCGTGA ATGATCGG TCATATGT
27	ACAATGGT ATCTTTTG AAGCGCAC GTAAACCG TAATTCGG AGCCCGGC
28	ACCTGGTC CGTCATAC TAAAAGGT AGAAGTAA GTAGGGTA CCGCTGGG
29	GCTCTTAA TATCTTTC GGTTGCGG GAGTAAGG
30	GTTTAGTG ATCCGTCA ACCATTGT TATTGGCC
31	CACCAGAA GAGCCACT TAACCTGG TGCGATCA ATCCCGGA CACGTAAA
32	CCACAAAG TCTTGTAT AAAGACTG TGCGAGAT CTTTACGC GCTTGAAC
33	TTCCTTTT CAACCAAA AGATGTTC CCCACTGT TTATATTA CAAAAGAT
34	CCGAATTA CGGTTTAC CGAGATGT TTACTTGC ACTTGGGA TTAGGATC
35	CTTGCCTG GGCCAATA CTTTGTGG TCCACTAT
36	CAGTCCGA CCTAGGAG CGGTCTGG CTGTGAGG
37	GCGTAAAG ATCTCGCA TTCTGGTG CCAAAACT AATCATAC TCACGGTT
38	TTCAGCAG CCTTACTC CCGCAACC TTTACGTG TTTAGGGA ATACAAGA
39	GTGCTCTT GAGCGGTG TTAAGAGC CTCGCTTC AAGTCCGT GTCGGTAG
40	AAAACTTC ACATATGA CCGATCAT TCACGACA AATCTCCC GCTAGATG
41	GGGAGCTG GCCGGGCT ATTACTTT TGCCATAT
42	GTTGTTCC ATAGTGGA AAGAGCAC GTTAACTC
43	GTAGTTTA CCCAGCGG TACCCTAC TTACTTCT AGTGTAAG CAACCCAT
44	CGACTCAA CCTCACAG CCAGACCG CATCTAGC ATTACCTT CACCGCTC
45	GCCGACTC GTTCAAGC ACTATTAT CTGGCTAT
46	ACTGAGGC GAGTTAAC TTATGTTC TAACCTGC
47	TCATGTGG GATCCTAA TCCCAAGT GCAAGTAA AAAAGGAA TTATGAGA
48	GATCACAA ATATGGCA AAAGTAAT ATGGGTTG CACTAAAC TATAGCAA
49	GAACATAA TTGCTATA TCGGACTG TGGGCGTA CCTACACC ACAGATGC
50	ACGGACTT GAAGCGAG GAAGTTTT GCGACTCA TTGCTGAG TTCGACGC
51	ATATAGGT CTGACACC CAGCTCCC AGGTAAGT GTATTATC GCGTCAAT
52	GGTGTAGG TACGCCCA TAAACTAC TTTTCCAC GTTGCCCT CGACCGAT
53	GGAGGGTT CTACCGAC ATATTGCT TTGCCCTT
54	TGTTGAGG GCAGGTTA ACCTATAT AGCCATTT
55	AGCTTGGG AACCGTGA GTATGATT AGTTTTGG CTGCTGAA TGTCCGTG
56	CGACCATT TGACACCC GAGTCGGC CTGACAGA CATCTTTT GGGGTCGT
57	AGCCTACG ATAGCCAG ATAATAGT TCTCATAA CAGGCAAG GGTGTCAG
58	GATAATAC ACTTACCT CCACATGA AGGTCTAC CAGAAACT ATGTACCG
59	GCGGAGAG AAATGGCT AATGGTCG ACAATATG
60	CGTCCAGC GCATCTGT GATCCACG TCAGATTT
61	AAAAGATG TCTGTCAG CCCAAGCT TAGCATAA AACTGTTG TGTCCTAT
62	GTCCAAGT AAGGGCAA AGCAATAT CACGGACA GGAACAAC GGGTGTCA
63	GCGTATGC TGAAATGG AACCCTCC TACGTTTT GCGATGGT GATGGGAA
64	TTTAGTTT GCGTCGAA CTCAGCAA TGAGTCGC TTGAGTCG GTGGATAA
65	ATCGCTCA ATTGACGC GACTTTGT GTAATCTC
66	CATACACA CATATTGT GCATACGC CCTGGTTT
67	CTGAGCGT ATCGGTCG AGGGCAAC GTGGAAAA TTGTGATC CCCTCCCA
68	AATAGGCT ATTCAAGA GCTGGACG TGCGCGAG GCACGTAT GTCTTGAG
69	TAGGGCTA AAATCTGA CGTGGATC TTATCCAC GCCTCAGT CCATTTCA
70	ACCATCGC AAAACGTA AAACTAAA TGACGTTT GAGGTGAT ACGTAAAG
71	GCGGCTGG ACGACCCC GCCGTGCG ACGACTGA
72	GTAATGCG AAACCAGG AGCCTATT CATCAAGG
73	GATATCCA CGGTACAT AGTTTCTG GTAGACCT CGTAGGCT AGAGCGGT
74	TATTAAGT TTTCGCCG TGAGCGAT CAATGACC TCATATGG ACACACCA
75	GATCCTGA GAGATTAC ACAAAGTC TGGGAGGG CCTCAACA TCTTGAAT
76	ATACGTGC CTCGCGCA ACGCTCAG CCGGTCTG TAGGGTGA CGTATTAG
77	CAAATCCC TTCCCATC AAATACGA CTAAGCCG
78	CGTTGCAC CCTTGATG ACTTAATA TTCAGGCT
79	CGTTTGTC ATAGGACA CAACAGTT TTATGCTA ACTTGGAC CGGGCATC
80	GCAGGGTG AAAGTGAG CCAGCCGC GCGCAGGG CTCGACGG TTAATTTG
81	ACTCTCGT TCAGTCGT CGCACGGC ACCGCTCT CTCTCCGC CGGCGAAA
82	CCATATGA GGTCATTG TGGATATC TCCGTGCA GTAAACGA TAGGTTTC

Table continues ...

ID	Sequence
83	CCAGTAAT AGCCTGAA CACCCTGC ATCGACGG
84	GCTAGCAT CTCAAGAC CTGGGAAT TAGGTAAG
85	CCGTCGAG CCCTGCGC GACAAACG CATATCCT GTTTACTT AGGGATGG
86	TCTGGCGC CGGCTTAG TCGTATTT GATGCCCG TGTGTATG CTCACTTT
87	GAAAGATC ATAACGTT GGGATTTG TTCCGCCT ATAGGTAA CTGGGTTT
88	GTCAATTT CTTTACGT ATCACCTC AAACGTCA TAGCCCTA CGGATCAA
89	GAATTATA TGGTGTGT CCATATCA ACAAACCC
90	GCGGCGAA CCGTCGAT GATCTTTC GGACCGAG
91	GGCTTGCA CTAATACG TCACCCTA CAGACCGG TCAGGATC ATTTTGGA
92	TTTGTAAA CTTACCTA ATTCCCAG TTGATCCG CGCATTAC AACGTTAT
93	CGGCCCGC CAAATTAA TAGTCGAT TACGCTTC
94	AGCTCATT CTCGGTCC AATTCGAG CATGTCCC
95	AAGCAAGA GAAACCTA TCGTTTAC TGCACGGA ACGAGAGT GCTGGCAT
96	TGGGCCCG GGGTTTGT TGATATGG TCCAAAAT GTGCAACG GGAAACAC
97	CTCGAATT GTGTTTCC ATGCTAGC GCCACACG ACTACTTA AGACTTAG
98	TTACCTAT AGGCGGAA AAATTGAC ATTTCTAC AATCGTGA CGGTTTCT
99	CTTCTACC TCTTAATG TATAATTC ATCCTGTG GATTACGG CGGCTTAG
100	TAAGTAGT CGTGTGGC TGCAAGCC CTCACCTG ACCTCTTA GCACAATA
101	CCTAACAT AAACCCAG CGTGACGG TGAATCTT
102	GTGTAAGA GGGACATG GGTAGAAG GAACTCTG
103	ATTACCGT CCATCCCT AAGTAAAC AGGATATG GCGCCAGA CCACAATG
104	GCCGATGG AAAACGAC GCGGGCCG GGAAGGAC CTTTAGGC CCAAGGAG
105	TCGCAACG GAAGCGTA ATCGACTA ATGCCAGC ATTACTGG CATTAAGA
106	CCGTAATC CACAGGAT TCTTGCTT ATCTTAGT CCCCAAGT GCTACATG
107	CATGGAAG CAGAGTTC CCATCGGC AAAATTGA
108	AACCGCGC CTAAGTCT TTAGAATC ACCTTAGG
109	GCCTAAAG GTCCTTCC ACGGTAAT CGGGGGTG TACGAACT CAACGCAC
110	CTCAGTTT AAGATTCA CCGTCACG CATTGTGG TTCGCCGC GTCGTTTT
111	CGACCTCT AGATCCCA ATGTTAGG AGCTAACA GATATTTC AAGGACAC
112	TCTTCAAA AGAAACCG TCACGATT GTAGAAAT TTTACAAA GACGGACA
113	AATGTTCT CTAAGCCG AGTAAACA ATCAAGGA
114	GAAGAGGT TCAATTTT AGAGGTCG AGGCGAGT
115	GTAGGCTC TATTGTGC TAAGAGGT CAGGTGAG CGGGCCCA ATTGTACC
116	ACGGAAAC TCCTACCC GCGCGGTT GATGTGCC TGCGAAAA GCAAGCCG
117	GTGGCGGA CCTAAGGT GATTCTAA TGTCCGTC AATGAGCT TGGGATCT
118	GAAATATC TGTTAGCT TTTGAAGA AGGCTCCG TCATTCCG TCCCCCAA
119	ACTCCGGC CTCCTTGG TATATTAG CTTACCCA
120	GACCCGAC ACTCGCCT GTTTCCGT GCGTTCTT
121	GCGATTCG CATGTAGC ACTTGGGG ACTAAGAT CGTTGCGA ATTTTCGC
122	GTTACCGC TATAATTA AGAACATT CGGGTTCG GGGAATTG AGTACAAG
123	GCGCAGTT TCCTTGAT TGTTTACT GGTACAAT TCTTACAC GGGTAGGA
124	TTTTCGCA GGCACATC GAGCCTAC CCGTCTAT AGCAGAAT GGTATAAA
125	CACCTACC GTGTCCTT CGTCCTGT CCCGTTCC
126	ATATTACA AAGAACGC GCGGTAAC ATCACATA
127	CGAGTCGT GTGCGTTG AGTTCGTA CACCCCCG AAACTGAG GAGAACAT
128	CCTAGGCG TAAGACCT GCCGGAGT AGACGTGC CACGCGAG CTATGTAG
129	CGTCGTCT TGGGTAAG CTAATATA GCGAAAAT CTTCCATG TAATTATA
130	CAATTCCC CGAACCCG CGAATCGC AGGACGAT CTCGTTCG GGTGCTAC
131	ACGTGGTC TATGTGAT CGCCTAGG CTAACCTC
132	CGGACACA CGGCTTGC TGCGACCA AGTTTTAT
133	CTCGCGTG GCACGTCT ACGACTCG TCGAACGA GCTCGGGT AGTCTCAA
134	ATTTTGGC GGAACGGG ACAGGACG ATGTTCTC ACCTCTTC AGGTCTTA
135	GCCAGACT TTTACGGC GGTAGGTG TAAGACCT CGCCCTGA CGTAGAAC
136	TAGCCGCC TTGGGGGA CGGAATGA CGGAGCCT TCCGCCAC CCCTGACA
137	GAGCTCAC CTTGTACT CTTATCCT TCACCAGG
138	GCGGATGC GAGGTTAG AGTCTGGC AACGGGTT
139	TCCCAATC TTTATACC ATTCTGCT ATAGACGG AACTGCGC GTCACTGC
140	AGGGATTG AGTACCCA TGTGTCCG AGATGGCA TATGACAG GTGAGCAC
141	TTCCGCGC ATAAAACT TGGTCGCA TGTCAGGG GTCGGGTC GCCGTAAA
142	TCAGGGCG AGGTCTTA GGCGGCTA CGTCTATT TAACGTGC CTGATCAC
143	GGACTGTT CTACATAG TGACTTGG ACGAGGTT
144	TCAGCGGC AACCCGTT CAATCCCT AGCCGTTC
145	GCATGCCG GTAGCACC CGAACGAG ATCGTCCT AGACGACG GATCTCCT
146	ATGACTCC AGGAGAGG GTGAGCTC CTACGTGG TGACGAAG AATACCGT
147	ATAACCAT CCTGGTGA AGGATAAG GCAGTGAC TGTAATAT TGGGTACT
148	CTGTCATA TGCCATCT GATTGGGA GAGTCCAG GCAATAAA GAAACTGC
149	TTCATTAC GTTCTACG CGATGCTT TTGCCACA
150	CCTACGCC GAACGGCT GGAGTCAT CTCGGCAT
151	AGTTTCCT TTGAGACT ACCCGAGC TCGTTCGA GCCAAAAT CAAGGACG
152	AGTTAGAT CCATGCGA AACAGTCC GTATAGCG CCCAGTGA GCTCGACA
153	GCGTTAAA AACCTCGT CCAAGTCA AGGAGATC GACCACGT CCTCTCCT
154	CTTCGTCA CCACGTAG CGGCATGC AGCACGCA TATTTGAC GTCTTGCG
155	AAGTGTTA ATGCCGAG ATCTAACT TGATTTTT
156	CCCGACCT GTGCTCAC GCACGAGA AGGCTTCC
157	TCACTGGG CGCTATAC AGGAAACT CGAGGGGC GAAGATAG AGAGCATA
158	TTAAGATG TGTGGCAA AAGCATCG CGTCCTTG GCATCCGC TCGCATGG
159	AAGAATGG TCTACATC GTAATGAA GGATAGAT AAGCAGAG CATTTTGA
160	TTCAGCGG GTGATCAG GCACGTTA AATAGACG GCGCGGAA GACCTCCG
161	AAGAGCAA ACGGTATT ATCCCGTT CACAGCGC
162	ACATGAGG AAAAATCA CCATTCTT GTGCCAGT
163	TTACCCTC GCAGTTTC TTTATTGC CTGGACTC ATGGTTAT CCCTTCCC
164	TTACCAGC GTGCGGCC AGGTCGGG ACAGCAGT ACCCATGG CGCACCAC
165	TATTTCCT GGAAGCCT TCTCGTGC CGGAGGTC GCCGCTGA GATGTAGA
166	CTCTGCTT ATCTATCC CCGCTGAA AGGATTAT TCAATAAT GCACCTGC
167	TCCCTGTC TGTCGAGC CCAGCGAC CACTTCTG
168	TTCTAGCC ACTGGCAC GCTGGTAA TCTGGGCT
169	CAGCTTGA CGCAAGAC GTCAAATA TGCGTGCT TTTAACGC TTTAGACC
170	ATTCAGAA TCGAGACC TTGCTCTT GGACGTCT CCATCAGA TGGAGCTG

Table continues ...

ID	Sequence
171	TGGACCAC GCGCTGTG AACGGGAT GGGAAGGG GGCGTAGG GGCCGCAC
172	CCATGGGT ACTGCTGT GAGGGTAA CACGTTGG ACTCCTAC TGCCCGCC
173	TTCAACAA TCAAAATG ATATTGTA GCGAGATA
174	CTACATTT AGCCCAGA TTCTGAAT CGCGGGCT
175	AACTACTG TATGCTCT CTATCTTC GCCCCTCG CATCTTAA ATATAACT
176	CTAGTAAC CGACCGTT GACAGGGA CCACCGCA TCTTTCTC GTATATCA
177	CTAGGACC CAGAAGTG GTCGCTGG GGTCTAAA TAACACTT GGTCTCGA
178	TCTGATGG AGACGTCC TCAAGCTG CTGCCCAG TGATTAGC GCTTAGGG
179	CATTCTAT AGCCCGCG GTTACTAG GCGCCTTG
180	GATGATCT GTGGTGCG CACAATTT GCTGGAAG
181	GAGAAAGA TGCGGTGG CAGTAGTT GTGATTAG AATTCGCT AAACCGGA
182	TATCCGCC TATCTCGC TACAATAT AGTTATAT CCTCATGT AACGGTCG
183	ATTAAGCT GTGGAATG TTGTTGAA CATACTCC TCATTAGG GTCCACCG
184	TTCCTAAC GCAGGTGC ATTATTGA ATAATCCT AGGAAATA CCACTGCC
185	AGGGTTCA CAGCTCCA AAGCCTGT CGAGGCGT
186	GACTCTCT CAAGGCGC AGCTTAAT TAGTAAGG
187	TGTTTTAA GGCGGGCA GTAGGAGT CCAACGTG GTGGTCCA CGGCGACG
188	CTTTTCTG CTTCCAGC AAATTGTG GGCAGTGG GGCTAGAA CATTCCAC
189	CTTGGGGG TGATATAC CTCACTAA ACTCCTTC
190	GTCCTATC CCTTACTA GCGGTACA CACCAAGA
191	AACCTACC CCCTAAGC GCTAATCA CTGGGCAG GGTCCTAG GTTGCAGA
192	TTACAGCG ACGCCTCG ACAGGCTT CGTCGCCG AAATGTAG CCATGAAT
193	TGTACCGC ATTCATGG AGATCATC CATTGACG TCGTTACG TTTTTTTT
194	CCTAATGA GGAGTATG GTTAGGAA TACCTGCT CCTGAATG TTTTTTTT
195	CTCCTACG CTCAATAG TGAACCCT TACTATCG
196	AGAGTTCT ACCAGCCA AGAGAGTC GCAGCCCT
197	CGTAACGA CGTCAATG TTAAAACA ACAACCGA
198	CATTCAGG AGCAGGTA CAGAAAAG GAACCGAC
199	TTTTTTTT CGATAGTA GGTAGGTT TTTTTTTT
200	TTTTTTTT AGGGCTGC CCCCCAAG TTTTTTTT
201	TTTTTTTT CGGTGGAC AGAACTCT TTTTTTTT
202	TTTTTTTT TCTTGGTG CGTAGGAG TTTTTTTT
203	TTTTTTTT TCGGTTGT CGCTGTAA TTTTTTTT
204	TTTTTTTT GTCGGTTC GATAGGAC TTTTTTTT
205	TTTTTTTT TCCGGTTT AGCGAATT CTAATCAC GGCGGATA TGGCTGGT
206	TTTTTTTT GAAGGAGT TTAGTGAG TCTGCAAC ATAGAATG CTATTGAG
207	CAGACTCT TTTTTTTT TTTTTTTT GAGCCGAT
208	CAGGGTAC TTTTTTTT TTTTTTTT GGACGAGC
209	TCTGATAT TTTTTTTT TTTTTTTT GCACTGCC
210	CGCTGATC TTTTTTTT TTTTTTTT GGGGATGA
211	CAAGGGAT ATCGGCTC CTTCGGAT TCTCTGGC
212	TACATCTC GCTCGTCC TCACCATA ATCCACCA
213	AGAGTAGG GGCAGTGC TCAGCCAA ACAGTGGG
214	GGAAGGAT TCATCCCC TTTAAGAA TGATCGCA
215	TATCATTG GCCAGAGA TGCTAACT GTATGACG
216	TGCATGGG TGGTGGAT CGCGAAGC TGTCGTGA
217	AGATGTTC CCCACTGT TTATATTA CAAAAGAT
218	TAACCTGG TGCGATCA ATCCCGGA CACGTAAA
219	ACCTGGTC CGTCATAC TAAAAGGT AGAAGTAA
220	CCGATCAT TCACGACA AATCTCCC GCTAGATG
221	ACAATGGT ATCTTTTG AAGCGCAC GTAAACCG
222	CCGCAACC TTTACGTG TTTAGGGA ATACAAGA
223	TACCCTAC TTACTTCT AGTGTAAG CAACCCAT
224	CCAGACCG CATCTAGC ATTACCTT CACCGCTC
225	CCGAATTA CGGTTTAC CGAGATGT TTACTTGC
226	CCACAAAG TCTTGTAT AAAGACTG TGCGAGAT
227	AAAGTAAT ATGGGTTG CACTAAAC TATAGCAA
228	GTGCTCTT GAGCGGTG TTAAGAGC CTCGCTTC
229	TCCCAAGT GCAAGTAA AAAAGGAA TTATGAGA
230	GCGTAAAG ATCTCGCA TTCTGGTG CCAAAACT
231	GAACATAA TTGCTATA TCGGACTG TGGGCGTA
232	ACGGACTT GAAGCGAG GAAGTTTT GCGACTCA
233	ATAATAGT TCTCATAA CAGGCAAG GGTGTCAG
234	GTATGATT AGTTTTGG CTGCTGAA TGTCCGTG
235	GGTGTAGG TACGCCCA TAAACTAC TTTTCCAC
236	CTCAGCAA TGAGTCGC TTGAGTCG GTGGATAA
237	ATATAGGT CTGACACC CAGCTCCC AGGTAAGT
238	AGCAATAT CACGGACA GGAACAAC GGGTGTCA
239	AGGGCAAC GTGGAAAA TTGTGATC CCCTCCCA
240	CGTGGATC TTATCCAC GCCTCAGT CCATTTCA
241	GATAATAC ACTTACCT CCACATGA AGGTCTAC
242	CGACCATT TGACACCC GAGTCGGC CTGACAGA
243	ACAAAGTC TGGGAGGG CCTCAACA TCTTGAAT
244	GCGTATGC TGAAATGG AACCCTCC TACGTTTT
245	AGTTTCTG GTAGACCT CGTAGGCT AGAGCGGT
246	AAAAGATG TCTGTCAG CCCAAGCT TAGCATAA
247	AATAGGCT ATTCAAGA GCTGGACG TGCGCGAG
248	ACCATCGC AAAACGTA AAACTAAA TGACGTTT
249	CGCACGGC ACCGCTCT CTCTCCGC CGGCGAAA
250	CAACAGTT TTATGCTA ACTTGGAC CGGGCATC
251	ATACGTGC CTCGCGCA ACGCTCAG CCGGTCTG
252	ATCACCTC AAACGTCA TAGCCCTA CGGATCAA
253	TATTAAGT TTTCGCCG TGAGCGAT CAATGACC
254	TCGTATTT GATGCCCG TGTGTATG CTCACTTT
255	TCACCCTA CAGACCGG TCAGGATC ATTTTGGA
256	ATTCCCAG TTGATCCG CGCATTAC AACGTTAT
257	CCATATGA GGTCATTG TGGATATC TCCGTGCA
258	GCAGGGTG AAAGTGAG CCAGCCGC GCGCAGGG

Table continues ...

ID	Sequence
259	TGATATGG TCCAAAAT GTGCAACG GGAAACAC
260	GAAAGATC ATAACGTT GGGATTTG TTCCGCCT
261	TCGTTTAC TGCACGGA ACGAGAGT GCTGGCAT
262	CCGTCGAG CCCTGCGC GACAAACG CATATCCT
263	CTCGAATT GTGTTTCC ATGCTAGC GCCACACG
264	TTACCTAT AGGCGGAA AAATTGAC ATTTCTAC
265	ATCGACTA ATGCCAGC ATTACTGG CATTAAGA
266	AAGTAAAC AGGATATG GCGCCAGA CCACAATG
267	TAAGTAGT CGTGTGGC TGCAAGCC CTCACCTG
268	TCACGATT GTAGAAAT TTTACAAA GACGGACA
269	CTTCTACC TCTTAATG TATAATTC ATCCTGTG
270	CCGTCACG CATTGTGG TTCGCCGC GTCGTTTT
271	TAAGAGGT CAGGTGAG CGGGCCCA ATTGTACC
272	GATTCTAA TGTCCGTC AATGAGCT TGGGATCT
273	CCGTAATC CACAGGAT TCTTGCTT ATCTTAGT
274	GCCGATGG AAAACGAC GCGGGCCG GGAAGGAC
275	TGTTTACT GGTACAAT TCTTACAC GGGTAGGA
276	CGACCTCT AGATCCCA ATGTTAGG AGCTAACA
277	ACTTGGGG ACTAAGAT CGTTGCGA ATTTTCGC
278	GCCTAAAG GTCCTTCC ACGGTAAT CGGGGGTG
279	ACGGAAAC TCCTACCC GCGCGGTT GATGTGCC
280	GAAATATC TGTTAGCT TTTGAAGA AGGCTCCG
281	CTAATATA GCGAAAAT CTTCCATG TAATTATA
282	AGTTCGTA CACCCCCG AAACTGAG GAGAACAT
283	TTTTCGCA GGCACATC GAGCCTAC CCGTCTAT
284	CGGAATGA CGGAGCCT TCCGCCAC CCCTGACA
285	GTTACCGC TATAATTA AGAACATT CGGGTTCG
286	ACAGGACG ATGTTCTC ACCTCTTC AGGTCTTA
287	ATTCTGCT ATAGACGG AACTGCGC GTCACTGC
288	TGGTCGCA TGTCAGGG GTCGGGTC GCCGTAAA
289	CAATTCCC CGAACCCG CGAATCGC AGGACGAT
290	CCTAGGCG TAAGACCT GCCGGAGT AGACGTGC
291	AGGATAAG GCAGTGAC TGTAATAT TGGGTACT
292	GCCAGACT TTTACGGC GGTAGGTG TAAGACCT
293	CGAACGAG ATCGTCCT AGACGACG GATCTCCT
294	CTCGCGTG GCACGTCT ACGACTCG TCGAACGA
295	AGGGATTG AGTACCCA TGTGTCCG AGATGGCA
296	TCAGGGCG AGGTCTTA GGCGGCTA CGTCTATT
297	CCAAGTCA AGGAGATC GACCACGT CCTCTCCT
298	ACCCGAGC TCGTTCGA GCCAAAAT CAAGGACG
299	CTGTCATA TGCCATCT GATTGGGA GAGTCCAG
300	GCACGTTA AATAGACG GCGCGGAA GACCTCCG
301	ATGACTCC AGGAGAGG GTGAGCTC CTACGTGG
302	AAGCATCG CGTCCTTG GCATCCGC TCGCATGG
303	TTTATTGC CTGGACTC ATGGTTAT CCCTTCCC
304	TCTCGTGC CGGAGGTC GCCGCTGA GATGTAGA
305	CTTCGTCA CCACGTAG CGGCATGC AGCACGCA
306	AGTTAGAT CCATGCGA AACAGTCC GTATAGCG
307	AACGGGAT GGGAAGGG GGCGTAGG GGCCGCAC
308	AAGAATGG TCTACATC GTAATGAA GGATAGAT
309	GTCAAATA TGCGTGCT TTTAACGC TTTAGACC
310	TCACTGGG CGCTATAC AGGAAACT CGAGGGGC
311	TTACCAGC GTGCGGCC AGGTCGGG ACAGCAGT
312	CTCTGCTT ATCTATCC CCGCTGAA AGGATTAT
313	GTCGCTGG GGTCTAAA TAACACTT GGTCTCGA
314	CTATCTTC GCCCCTCG CATCTTAA ATATAACT
315	CCATGGGT ACTGCTGT GAGGGTAA CACGTTGG
316	ATTATTGA ATAATCCT AGGAAATA CCACTGCC
317	ATTCAGAA TCGAGACC TTGCTCTT GGACGTCT
318	TACAATAT AGTTATAT CCTCATGT AACGGTCG
319	GTAGGAGT CCAACGTG GTGGTCCA CGGCGACG
320	AAATTGTG GGCAGTGG GGCTAGAA CATTCCAC
321	TCTGATGG AGACGTCC TCAAGCTG CTGCCCAG
322	CTAGTAAC CGACCGTT GACAGGGA CCACCGCA
323	ACAGGCTT CGTCGCCG AAATGTAG CCATGAAT
324	ATTAAGCT GTGGAATG TTGTTGAA CATACTCC
325	GCTAATCA CTGGGCAG GGTCCTAG GTTGCAGA
326	GAGAAAGA TGCGGTGG CAGTAGTT GTGATTAG
327	TGTACCGC ATTCATGG AGATCATC CATTGACG
328	CCTAATGA GGAGTATG GTTAGGAA TACCTGCT
329	TTAGTGAG TCTGCAAC ATAGAATG CTATTGAG
330	AGCGAATT CTAATCAC GGCGGATA TGGCTGGT
331	ACCTTTTA ATCATCTC
332	GGGAGATT TAGCGGTC
333	CCTACTCT CACCTTTC
334	ATCCTTCC AACATCTC
335	CTTACACT TTCGGGAG
336	CCCATGCA GAAAGATA
337	ACATCTCG GAAAGGTG
338	CCAGGTTA AGTGGCTC
339	GACCAGGT CTCCTAGG
340	ATGATCGG TCATATGT
341	TTCCTTTT CAACCAAA
342	CACCAGAA GAGCCACT
343	GTAGGGTA CCGCTGGG
344	AAAACTTC ACATATGA
345	TAATTCGG AGCCCGGC
346	TTCAGCAG CCTTACTC

Table continues ...

ID	Sequence
347	GTAGTTTA CCCAGCGG
348	CGACTCAA CCTCACAG
349	ACTTGGGA TTAGGATC
350	CTTTACGC GCTTGAAC
351	GATCACAA ATATGGCA
352	AAGTCCGT GTCGGTAG
353	TCATGTGG GATCCTAA
354	AATCATAC TCACGGTT
355	CCTACACC ACAGATGC
356	TTGCTGAG TTCGACGC
357	AGCCTACG ATAGCCAG
358	AGCTTGGG AACCGTGA
359	GTTGCCCT CGACCGAT
360	TTTAGTTT GCGTCGAA
361	GTATTATC GCGTCAAT
362	GTCCAAGT AAGGGCAA
363	CTGAGCGT ATCGGTCG
364	TAGGGCTA AAATCTGA
365	CAGAAACT ATGTACCG
366	CATCTTTT GGGGTCGT
367	GATCCTGA GAGATTAC
368	GCGATGGT GATGGGAA
369	GATATCCA CGGTACAT
370	AACTGTTG TGTCCTAT
371	GCACGTAT GTCTTGAG
372	GAGGTGAT ACGTAAAG
373	ACTCTCGT TCAGTCGT
374	CGTTTGTC ATAGGACA
375	TAGGGTGA CGTATTAG
376	GTCAATTT CTTTACGT
377	TCATATGG ACACACCA
378	TCTGGCGC CGGCTTAG
379	GGCTTGCA CTAATACG
380	TTTGTAAA CTTACCTA
381	GTAAACGA TAGGTTTC
382	CTCGACGG TTAATTTG
383	TGGGCCCG GGGTTTGT
384	ATAGGTAA CTGGGTTT
385	AAGCAAGA GAAACCTA
386	GTTTACTT AGGGATGG
387	ACTACTTA AGACTTAG
388	AATCGTGA CGGTTTCT
389	TCGCAACG GAAGCGTA
390	ATTACCGT CCATCCCT
391	ACCTCTTA GCACAATA
392	TCTTCAAA AGAAACCG
393	GATTACGG CGGCTTAG
394	CTCAGTTT AAGATTCA
395	GTAGGCTC TATTGTGC
396	GTGGCGGA CCTAAGGT
397	CCCCAAGT GCTACATG
398	CTTTAGGC CCAAGGAG
399	GCGCAGTT TCCTTGAT
400	GATATTTC AAGGACAC
401	GCGATTCG CATGTAGC
402	TACGAACT CAACGCAC
403	TGCGAAAA GCAAGCCG
404	TCATTCCG TCCCCCAA
405	CGTCGTCT TGGGTAAG
406	CGAGTCGT GTGCGTTG
407	AGCAGAAT GGTATAAA
408	TAGCCGCC TTGGGGGA
409	GGGAATTG AGTACAAG
410	ATTTTGGC GGAACGGG
411	TCCCAATC TTTATACC
412	TTCCGCGC ATAAAACT
413	CTCGTTCG GGTGCTAC
414	CACGCGAG CTATGTAG
415	ATAACCAT CCTGGTGA
416	CGCCCTGA CGTAGAAC
417	GCATGCCG GTAGCACC
418	GCTCGGGT AGTCTCAA
419	TATGACAG GTGAGCAC
420	TAACGTGC CTGATCAC
421	GCGTTAAA AACCTCGT
422	AGTTTCCT TTGAGACT
423	GCAATAAA GAAACTGC
424	TTCAGCGG GTGATCAG
425	TGACGAAG AATACCGT
426	TTAAGATG TGTGGCAA
427	TTACCCTC GCAGTTTC
428	TATTTCCT GGAAGCCT
429	TATTTGAC GTCTTGCG
430	CCCAGTGA GCTCGACA
431	TGGACCAC GCGCTGTG
432	AAGCAGAG CATTTTGA
433	CAGCTTGA CGCAAGAC
434	GAAGATAG AGAGCATA

Table continues ...

ID	Sequence
435	ACCCATGG CGCACCAC
436	TCAATAAT GCACCTGC
437	CTAGGACC CAGAAGTG
438	AACTACTG TATGCTCT
439	ACTCCTAC TGCCCGCC
440	TTCCTAAC GCAGGTGC
441	CCATCAGA TGGAGCTG
442	TATCCGCC TATCTCGC
443	TGTTTTAA GGCGGGCA
444	CTTTTCTG CTTCCAGC
445	TGATTAGC GCTTAGGG
446	TCTTTCTC GTATATCA
447	TTACAGCG ACGCCTCG
448	TCATTAGG GTCCACCG
449	AACCTACC CCCTAAGC
450	AATTCGCT AAACCGGA
451	TCGTTACG TTTTTTTT
452	CCTGAATG TTTTTTTT
453	TTTTTTTT GAAGGAGT
454	TTTTTTTT TCCGGTTT

The following bricks were used in all structures investigated:

1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 17, 18, 21, 22, 29, 30, 35, 36, 41, 42, 45, 46, 53, 54, 59, 60, 65, 66, 71, 72, 77, 78, 83, 84, 89, 90, 93, 94, 101, 102, 107, 108, 113, 114, 119, 120, 125, 126, 131, 132, 137, 138, 143, 144, 149, 150, 155, 156, 161, 162, 167, 168, 173, 174, 179, 180, 185, 186, 189, 190, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204

In addition to the bricks common to all structures, the following bricks were used for each class of structure studied.

The **no-BB** system (330 bricks in total):

207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454

The **edge-BB** system (268 bricks in total):

8, 15, 19, 23, 26, 28, 31, 34, 37, 40, 43, 47, 50, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 95, 98, 100, 103, 106, 109, 112, 115, 118, 121, 124, 127, 130, 133, 136, 139, 142, 145, 148, 151, 154, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187, 191, 194, 205, 208, 210, 211, 212, 215, 217, 221, 222, 224, 226, 227, 228, 231, 233, 237, 238, 240, 242, 243, 244, 247, 249, 253, 254, 256, 258, 259, 260, 263, 265, 269, 270, 272, 274, 275, 276, 279, 281, 285, 286, 288, 290, 291, 292, 295, 297, 301, 302, 304, 306, 307, 308, 311, 313, 317, 318, 320, 322, 323, 324, 327, 329, 332, 334, 335, 336, 339, 341, 345, 346, 348, 350, 351, 352, 355, 357, 361, 362, 364, 366, 367, 368, 371, 373, 377, 378, 380, 382, 383, 384, 387, 389, 393, 394, 396, 398, 399, 400, 403, 405, 409, 410, 412, 414, 415, 416, 419, 421, 425, 426, 428, 430, 431, 432, 435, 437, 441, 442, 444, 446, 447, 448, 451, 453

The **face-BB** system (268 bricks in total):

7, 16, 20, 24, 25, 27, 32, 33, 38, 39, 44, 48, 49, 51, 56, 57, 62, 63, 68, 69, 74, 75, 80, 81, 86, 87, 92, 96, 97, 99, 104, 105, 110, 111, 116, 117, 122, 123, 128, 129, 134, 135, 140, 141, 146, 147, 152, 153, 158, 159, 164, 165, 170, 171, 176, 177, 182, 183, 188, 192, 193, 206, 207, 209, 213, 214, 216, 218, 219, 220, 223, 225, 229, 230, 232, 234, 235, 236, 239, 241, 245, 246, 248, 250, 251, 252, 255, 257, 261, 262, 264, 266, 267, 268, 271, 273, 277, 278, 280, 282, 283, 284, 287, 289, 293, 294, 296, 298, 299, 300, 303, 305, 309, 310, 312, 314, 315, 316, 319, 321, 325, 326, 328, 330, 331, 333, 337, 338, 340, 342, 343, 344, 347, 349, 353, 354, 356, 358, 359, 360, 363, 365, 369, 370, 372, 374, 375, 376, 379, 381, 385, 386, 388, 390, 391, 392, 395, 397, 401, 402, 404, 406, 407, 408, 411, 413, 417, 418, 420, 422, 423, 424, 427, 429, 433, 434, 436, 438, 439, 440, 443, 445, 449, 450, 452, 454

The **half-face-BB** system (299 bricks in total):

7, 16, 20, 32, 38, 39, 44, 56, 62, 63, 69, 80, 86, 87, 92, 104, 110, 111, 117, 128, 134, 135, 141, 152, 158, 159, 165, 176, 182, 183, 188, 207, 209, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221, 223, 225, 227, 229, 230, 231, 232, 233, 234, 235, 236, 237, 239, 241, 243, 245, 246, 247, 248, 249, 250, 251, 252, 253, 255, 257, 259, 261, 262, 263, 264, 265, 266, 267, 268, 269, 271, 273, 275, 277, 278, 279, 280, 281, 282, 283, 284, 285, 287, 289, 291, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 305, 307, 309, 310, 311, 312, 313, 314, 315, 316, 317, 319, 321, 323, 325, 326, 327, 328, 329, 330, 331, 333, 335, 337, 338, 339, 340, 341, 342, 343, 344, 345, 347, 349, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 363, 365, 367, 369, 370, 371, 372, 373, 374, 375, 376, 377, 379, 381, 383, 385, 386, 387, 388, 389, 390, 391, 392, 393, 395, 397, 399, 401, 402, 403, 404, 405, 406, 407, 408, 409, 411, 413, 415, 417, 418, 419, 420, 421, 422, 423, 424, 425, 427, 429, 431, 433, 434, 435, 436, 437, 438, 439, 440, 441, 443, 445, 447, 449, 450, 451, 452, 453, 454

The **all-BB** system (206 bricks in total):

7, 8, 15, 16, 19, 20, 23, 24, 25, 26, 27, 28, 31, 32, 33, 34, 37, 38, 39, 40, 43, 44, 47, 48, 49, 50, 51, 52, 55, 56, 57, 58, 61, 62, 63, 64, 67, 68, 69, 70, 73, 74, 75, 76, 79, 80, 81, 82, 85, 86, 87, 88, 91, 92, 95, 96, 97, 98, 99, 100, 103, 104, 105, 106, 109, 110, 111, 112, 115, 116, 117, 118, 121, 122, 123, 124, 127, 128, 129, 130, 133, 134, 135, 136, 139, 140, 141, 142, 145, 146, 147, 148, 151, 152, 153, 154, 157, 158, 159, 160, 163, 164, 165, 166, 169, 170, 171, 172, 175, 176, 177, 178, 181, 182, 183, 184, 187, 188, 191, 192, 193, 194, 205, 206

6 REFERENCES

[1] Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, "Three-dimensional structures self-assembled from DNA bricks," Science **338**, 1177 (2012).

[2] M. Sajfutdinow, K. Uhlig, A. Prager, C. Schneider, B. Abel, and D. M. Smith, "Nanoscale patterning of self-assembled monolayer (SAM)-functionalised substrates with single molecule contact printing," Nanoscale **9**, 15098 (2017).

[3] M. I. Mitov, M. L. Greaser, and K. S. Campbell, "GelBandFitter – A computer program for analysis of closely spaced electrophoretic and immunoblotted bands," Electrophoresis **30**, 848 (2009).

[4] S. W. Provencher, "A constrained regularization method for inverting data represented by linear algebraic or integral equations," Comput. Phys. Commun. **27**, 213 (1982).

[5] S. Hansen, "DLSanalysis.org: a web interface for analysis of dynamic light scattering data," Eur. Biophys. J. **47**, 179 (2018).

[6] H. Zipper, H. Brunner, J. Bernhagen, and F. Vitzthum, "Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications," Nucleic Acids Res. **32**, e103 (2004).

[7] J.-P. J. Sobczak, T. G. Martin, T. Gerling, and H. Dietz, "Rapid folding of DNA into nanoscale shapes at constant temperature," Science **338**, 1458 (2012).

[8] J. SantaLucia Jr and D. Hicks, "The thermodynamics of DNA structural motifs," Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).

[9] R. T. Koehler and N. Peyret, "Thermodynamic properties of DNA sequences: Characteristic values for the human genome," Bioinformatics **21**, 3333 (2005).

[10] W. Wang, T. Lin, S. Zhang, T. Bai, Y. Mi, and B. Wei, "Self-assembly of fully addressable DNA nanostructures from double crossover tiles," Nucleic Acids Res. **44**, 7989 (2016).

[11] Y. Ke, L. L. Ong, W. Sun, J. Song, M. Dong, W. M. Shih, and P. Yin, "DNA brick crystals with prescribed depths," Nat. Chem. **6**, 994 (2014).

[12] S. Giglio, P. T. Monis, and C. P. Saint, "Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR," Nucleic Acids Res. **31**, e136 (2003).

[13] D. Suh and J. B. Chaires, "Criteria for the mode of binding of DNA binding agents," Bioorg. Med. Chem. **3**, 723 (1995).

[14] K. M. Ririe, R. P. Rasmussen, and C. T. Wittwer, "Product differentiation by analysis of DNA melting curves during the polymerase chain reaction," Anal. Biochem. **245**, 154 (1997).