Nat Comput
DOI 10.1007/s11047-014-9431-z

Exponential replication of patterns in the signal tile assembly model

Alexandra Keenan · Robert Schweller · Xingsi Zhong

Abstract Chemical self-replicators are of considerable interest in the field of nanomanufacturing and as a model for evolution. We introduce the problem of self-replication of rectangular two-dimensional patterns in the practically motivated signal tile assembly model (STAM) (Padilla et al. Asynchronous signal passing for tile self-assembly: fuel efficient computation and efficient assembly of shapes, 2013). The STAM is based on the tile assembly model (TAM) which is a mathematical model of self-assembly in which DNA tile monomers may attach to other DNA tile monomers in a programmable way. More abstractly, four-sided tiles are assigned glue types to each edge, and self-assembly occurs when singleton tiles bind to a growing assembly, if the glue types match and the glue binding strength exceeds some threshold. The signal tile extension of the TAM allows signals to be propagated across assemblies to activate glues or break apart assemblies. Here, we construct a pattern replicator that replicates a two-dimensional input pattern over some fixed alphabet of size ϕ with $O(\phi)$ tile types, $O(\phi)$ unique glues, and a signal complexity of $O(1)$. Furthermore, we show that this replication system displays exponential growth in n, the number of replicates of the initial patterned assembly.

A preliminary version of this work appeared in Keenan et al. (2013)

A. Keenan (✉) · R. Schweller · X. Zhong
Department of Computer Science, University of Texas - Pan American Engineering Building, 3rd Floor 1012 W. University Dr., Edinburg, TX 78541, USA
e-mail: alexandra.b.keenan@gmail.com

R. Schweller
e-mail: rtschweller@utpa.edu

X. Zhong
e-mail: zhongx@utpa.edu

Keywords Self-replication · Computational geometry · Self-assembly · Wang tiles · Tile self-assembly

1 Introduction

Artificial self-replicating systems have been the subject of various investigations since John von Neumann first outlined a detailed conceptual proposal for a non-biological self-replicating system (Marchal 1998).

Gunter von Kiedrowski, who demonstrated the first enzyme-free abiotic replication system in 1986 (von Kiedrowski 1986), describes a model that can be used to conceptualize template-directed self-replication (Patzke and von Kiedrowski 2007). In this model, minimal template-directed self-replicating systems consist of an autocatalytic template molecule, and two or more substrate molecules that bind the template molecule and join together to form another template molecule. To date, simple self-replicating systems have been demonstrated in the laboratory with nucleic acids, peptides, and other small organic molecules (Paul and Joyce 2002; von Kiedrowski 1986; Zielinski and Orgel 1987; Tjivikua et al. 1990; Lincoln and Joyce 2009).

Given that substrate molecules must come together without outside guidance to replicate the template, a template-directed self-replicating system is necessarily a self-assembling system. In theoretical computer science, the tile assembly model (TAM) has become the most commonly used model to describe various self-assembly processes (Winfree 1998). Many model varients have been described since Erik Winfree first introduced the TAM, however models that are most relevant to self-replicating systems are those that allow for assembly breakage. These include the enzyme staged assembly model (Abel et al. 2010), the temperature programming model (Kao and Schweller 2006), the signal tile assembly model (Padilla et al. 2013, 2011), and the use of negative glues (Reif et al. 2005).

A theoretical result, the replication of arbitrary 0-genus shapes (polygons with no "holes"), has been shown within the staged assembly model with the use of RNAse enzymes (Abel et al. 2010). Replication and evolution of combinatorial 'genomes' via crystal-like growth and breakage have also been demonstrated in the laboratory using DNA tile monomers (Schulman et al. 2012). Under this replication mechanism, a DNA crystal ribbon has a sequence of information, or genome, in each row. Upon mechanical breakage, the daughter crystal continues to grow and copy the genome of the mother crystal. It was further shown that the fidelity of the replication process is sufficiently high for Darwinian evolution. Such simple, enzyme-free systems are of particular importance to the study of the origins of life.

A template-directed method of exponential self-replication within the tile assembly model, where the child molecule autodissociates from and is identical to the parent (as is found in biological systems), has not yet been described. Here, we present a theoretical basis for template-directed exponential self-replication in the practically motivated signal tile assembly model (STAM). Specifically, we consider the problem of self-replication of rectangular two-dimensional patterns in the STAM. The STAM is a powerful model of tile self-assembly in which activation, via binding, of a glue on an individual tile may turn other glues either on or off elsewhere on the tile (Padilla et al. 2013). In this way, signals may be propagated across distances greater than a single tile and assemblies may be broken apart. DNA strand displacement reactions provide a plausible physical basis for the signaling cascades used in the STAM. DNA strand displacement occurs when two DNA strands with at least partial complementarity hybridize with each other, which can displace prehybridized strands. In the STAM, these reactions may be queued to result in a cascade that ultimately turns a glue "on" by releasing a prehybridized strand. Conversely these queued reactions could turn a glue "off" by binding a free strand, thus making it unavailable to interact with other glues.

An important objective of nanotechnology is to manufacture things inexpensively, thus the prospect of self-replicating materials with useful patterns or functions is enticing. An exponential rate of growth is desirable for low-cost manufacturing of nanoscale devices, and we approach this problem with the goal of exponential growth in mind.

1.1 Summary of results

The STAM of Padilla et al. (2013) is briefly defined formally in Sect. 2, followed by our formal definition of exponential replication.

We first present a 2D signal tile system that replicates a linear pattern and then extend this mechanism to present our main result in Sect. 4—there exists a single, general purpose 2D signal tile system that exponentially replicates any rectangular 2D pattern (Theorem 1).

2 Definitions

2.1 Basic definitions

2.1.1 Multisets

A multiset is an ordered pair (S, m) where S is a subset of some universe set U and m is a function from U to $\mathbb{N} \bigcup \{\infty\}$ with the property that $m(x) \geq 1$ for all $x \in S$ and $m(x) = 0$ for all $x \notin S$. A multiset models a collection of items in which there are a positive number of copies $m(x)$ of each element x in the collection (called the multiplicity of x). For a multi-set $A = (S, m)$ and $x \in S$, we will use notation $A(x) = m(x)$ to refer to the multiplicity of item x, and $|A| \triangleq \sum_{a \in S} m(a)$ to refer to the *size* of A. For multisets $B = (b, m)$ and $A = (a, n)$, define $B \bigcup A$ to be the multiset $(a \bigcup b, m')$ where $m'(x) = m(x) + n(x)$. If $m(x) \geq n(x)$ for all $x \in U$, then define $B - A$ to be the multiset $(b', m'(x))$ where $b' = \{x \in b | m(x) - n(x) \geq 1\}$ and $m'(x) = m(x) - n(x)$. We use standard set notation $\{a_1, \ldots, a_r\}$ to denote multi-sets with the multiplicity of an item a being inferred by the number of i such that $a_i = a$.

2.1.2 Patterns

Let ϕ be a set of labels that contains at least one particular label $null \in \phi$ which conceptually denotes a blank, non-existent label. Informally, a 2D pattern is defined to be a mapping of 2D coordinates to elements of ϕ. Further, as these patterns will denote patterns on the surface of free floating tile assemblies, we add that patterns are equal up to translation. Formally, a 2D pattern over set ϕ is any set $\{f_{\Delta_x, \Delta_y}(x, y) | \Delta_x, \Delta_y \in \mathbb{Z}\}$ where $f : \mathbb{Z}^2 \rightarrow \phi$, and $f_{\Delta_x, \Delta_y}(x, y) = f(x + \Delta_x, y + \Delta_y)$. In this paper we focus on the the class of *rectangular* patterns in which the `null` labels occur at all positions outside of a rectangular box, with positions within the box labeled arbitrarily with non `null` labels.

2.2 Signal tile model

In this section we define the STAM by defining the concepts of an *active tile* consisting of a unit square with *glue slots* along the faces of the tile, as well as *assemblies* which consist of a collection of active tiles positioned on the integer lattice. We further define a set of three *reactions* (*break*, *combination*, and *glue-flip* reactions) which define how a set of assemblies can change over time. Figure 1 represents each of these concepts pictorially to help clarify the following technical definitions. Please see Padilla et al. (2013) for a more detailed presentation of the STAM.

2.2.1 Glue slots

Glue slots are the signal tile equivalent of glues in the standard tile assembly model with the added functionality of being able to be in one of three states, *on*, *off*, or *latent*, as well as having a *queued command* of *on*, *off*, or -, denoting if the glue is queued to be turned on, turned off, or has not been queued to change state. Formally, we denote a glue slot as an ordered triple $(g, s, q) \in \Sigma \times \{on, off, latent\} \times \{on, off, -\}$ where Σ is some given set of labels referred to as the *glue type* alphabet. For a given glue slot $x = (g, s, q)$, we define the *type* of x to be g, the *state* of x to be s, and the *queued action* of x to be q.

2.2.2 Active tiles

An active tile is a 4-sided unit square with each edge having a sequence of *glue slots* $g_1, \ldots g_r$ for some positive integer r, as well as an additional *label* taken from a set of symbols ϕ. For simplicity of the model, we further require that the glue type of each g_i on each tile face is the same (although state and queued commands may be different), and that the glue type of g_i is distinct from the glue type of g_j if $i \neq j$. For an active tile t, let $t_{d,i}$ denote the glue slot g_i on face d of active tile t.

Finally, an active tile t has an associated *signal function* $f_t(d, i)$ which assigns to each glue slot i on each tile side d a corresponding set of triples consisting of a glue slot, a side, and a command, which together denote which glue slots of each tile face should be turned on or off in the event that slot i on face d becomes bonded. Formally, each active tile t has an associated *signal function* $f : \{north, south, east, west\} \times \{1, \ldots r\} \rightarrow \mathcal{P}(\{north, east, south, west\} \times \{1, \ldots r\} \times \{on, off\})$. For the remainder of this paper we will use the term *tile* and *active tile* interchangeably.

2.2.3 Assemblies

An assembly is a set of active tiles whose centers are located at integer coordinates, and no two tiles in the set are at the same location. For an assembly A, define the weighted graph $G_A = (V, E)$ such that $V = A$, and for any pair of tiles $a, b \in V$, the weight of edge (a, b) is defined to be 0 if a and b do not have an overlapping face, and if a and b have overlapping faces d_a and d_b, the weight is defined to be $|\{i : state(a_{d_a,i}) = state(b_{d_b,i}) = on\}|$. That is, the weight of two adjacent tiles is the total number of matching glue types from a and b's overlapping edges that are both in state on. Conceptually, each such pair of equal, on glues represents a bond between a and b and thus increases the bonding strength between the tiles by 1 unit. For a positive integer τ, an assembly A is said to be τ-*stable* if the min-cut of the bond graph G_A is at least τ. For an assembly A, there is an associated pattern $p(A)$ defined by mapping the labels of each tile to corresponding lattice positions, and mapping the *null* label to lattice positions corresponding to locations not covered by the assembly.

2.2.4 Reactions

A reaction is an ordered pair of multi-sets of assemblies. Conceptually, a reaction (A, B) represents the assemblies of multi-set A replacing themselves with the assemblies in multi-set B. For a reaction $r = (A, B)$, let r_{in} denote the multi-set A, and r_{out} denote the multi-set B. For a set of reactions R, let $R_{in} = \bigcup_{r \in R} r_{in}$ and $R_{out} = \bigcup_{r \in R} r_{out}$.

A reaction (A, B) is said to be *valid* for a given temperature τ if it is either a *break*, *combination*, or *glue-flip* reaction as defined below:

- *Break reaction* A reaction $(A = \{a\}, B = \{b_1, b_2\})$ with $|A| = 1$ and $|B| = 2$ is said to be a break reaction if the bond graph of a has a cut of strength less than τ that separates a into assemblies b_1 and b_2.
- *Combination reaction* A reaction $(A = \{a_1, a_2\}, B = \{b\})$ with $|A| = 2$ and $|B| = 1$ is said to be a combination reaction if a_1 and a_2 are *combinable* into assembly b (see definition below).
- *Glue-flip reaction* A reaction $(A = \{a\}, B = \{b\})$ with $|A| = 1$ and $|B| = 1$ is said to be a glue-flip reaction if assembly b can be obtained from assembly a by changing the state of a single glue slot x in b to either `on` from `latent` if x has queued command `on`, or `off` from `on` or `latent` if x has queued command `off`. Note that transitions among `latent`, `on`, and `off` form an acyclic graph with sink state `off`, implying glues states can be adjusted at most twice. This models the "fire once" property of signals.

Two assemblies a_1 and a_2 are said to be *combinable* if a_1 and a_2 can be translated such that a_1 and a_2 have no overlapping tile bodies, but have at least τ `on`, matching glues connecting tiles from a_1 to tiles from a_2. Given this translated pair of assemblies, consider the product assembly b to be the assemblies a_1 and a_2 merged with the queued commands for each glue slot set according to the specifications of the glue functions for each tile with newly bonded `on` glues along the cut between a_1 and a_2. In this case we say a_1 and a_2 are *combinable* into assembly b. See

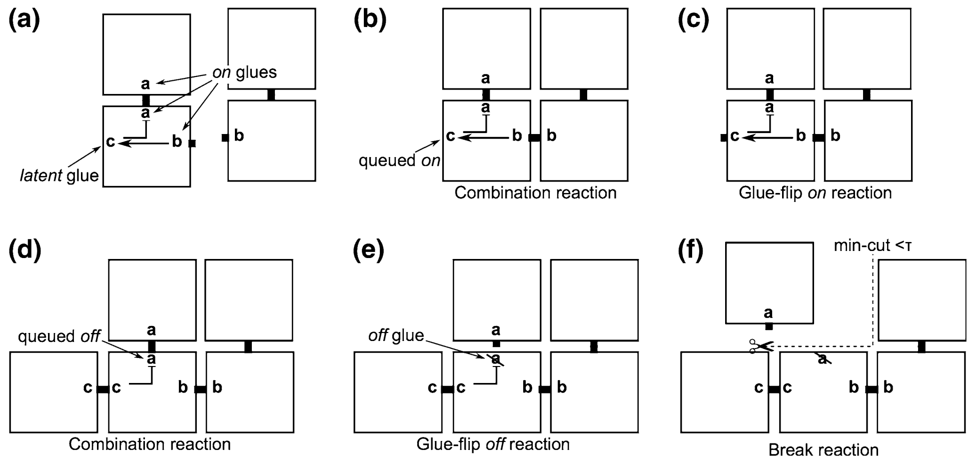

Fig. 1 This sequence (**a**–**f**) demonstrates the reaction types, glue states, and queued commands defined in the STAM

Fig. 1 for example reactions and Padilla et al. (2013) for a more detailed presentation of the model.

2.2.5 Batches

A batch is a multi-set of assemblies, i.e., a set of assemblies such that each assembly has a positive or infinite multiplicity. A batch B is said to be τ *-transitional* to a batch B' if the application of one of the break, combination, or transition rules at temperature τ can be applied to B to get B'. A *batch sequence* for some temperature τ is any sequence of batches $\langle a_1, \ldots a_r \rangle$ such that a_i is τ-transitional to a_{i+1} for each i from 1 to $r-1$.

2.2.6 Signal tile system

A signal tile system is an ordered pair (B, τ) where B is a batch referred to as the *initial seed* batch, and τ is a positive integer referred to as the temperature of the system. Any batch B' is said to be *producible* by (B, τ) if there exists a batch sequence $\langle B_1, \ldots, B_r \rangle$ with respect to temperature τ such that $B' = B_r$ and $B = B_1$, i.e., B' is reachable from B by a sequence of τ-transitions.

2.3 Exponential replication

Our first primary definition towards the concept of exponential replication defines a transition between batches in which multiple reactions may occur in parallel to complete the transition. By counting the number of such parallelized transitions we are able to define the number of time steps taken for one batch to transform into another, and in turn can define the concept of exponential replication.

However, to avoid reliance on highly unlikely reactions, we parameterize our definition with a positive integer c which dictates that any feasible combination reaction should involve at least one combinate with at least multiplicity c. By doing so, our exponential replication definition will be able to exclude systems that might rely on the highly unlikely combination of low concentration combinates (but will still consider such reactions in a worst-case scenario by requiring the subsequent monotonicity requirement). The following definition formalizes this concept.

Definition 1 ((τ, c)-*transitional distance*) We say a batch B is (τ, c)-transitional to a batch B', with notation $B \rightarrow_{\tau,c} B'$, if there exists a set of reactions $R = COMBO \bigcup BREAK \bigcup FLIP$, where COMBO, BREAK, and FLIP partition R into the combination, break, and flip type reactions, such that:

1. $B - R_{in}$ is defined and $B' = B - R_{in} + R_{out}$.
2. For each $(\{x, y\}, \{z\}) \in COMBO$, the multiplicity of either x or y in $B - R_{in}$ is at least c.

Further, we use notation $B \rightarrow^t_{\tau,c} B'$ if there exists a sequence $\langle B_1, \ldots, B_t \rangle$ such that $B_1 = B$, $B_t = B'$, and $B_i \rightarrow_{\tau,c} B_{i+1}$ for i from 1 to $t-1$. We define the (τ, c)-transitional distance from B to B' to be the smallest positive integer t such that $B \rightarrow^t_{\tau,c} B'$.

Our next primary concept used to define exponential replication is the concept of monotonicity which requires that a sequence of batches (regardless of how likely) has the property that each subsequent batch in the sequence is at least as close (in terms of (τ, c)-transition distance) to becoming an element of a given goal set of batches as any previous batch in the sequence.

Definition 2 (*Monotonicity*) Let B be a batch of assemblies, τ a positive integer, and G a set of (goal) batches. We say B grows monotonically towards G at temperature τ if for all batch sequences $\langle B, \ldots, B' \rangle$ at temperature τ such that $B \rightarrow^t_{\tau,c} g$ for some $g \in G$, then $B' \rightarrow^{t'}_{\tau,c} g'$ for some $g' \in G$ and $t' \leq t$.

Note that g' in the above definition may differ from g. This means that B is not required to grow steadily towards any particular element of G, but simply must make steady progress towards becoming an element of G.

We now apply the concepts of (τ, c)-transition distance and monotonicity to define exponential replication of patterns. Informally, an STAM system is said to replicate the pattern of an assembly a if it is always guaranteed to have a logarithmic (in n) (τ, c)-transitional distance that will create at least n copies of a shape with a's pattern for any integer n. Further, to ensure that the system makes steady progress towards the goal of n copies, we further require the property of *monotonicity* which states that the number of transitions needed to attain the goal of n copies never increases, regardless of the sequence of reactions.

Definition 3 (*Exponential Replication*) Let B_p^n denote the set of all batches which contain an n or higher multiplicity assembly with pattern p. A system $T = (B, \tau)$ exponentially replicates the pattern of assembly a if for all positive integers n and c:

1. $B \bigcup \{a\} \rightarrow_{\tau,c}^{t} B'$ for some $B' \in B_{p(a)}^n$ and $t = O(poly(|a|)\log(cn))$.
2. B grows monotonically towards $B_{p(a)}^n$.

Given the concept of a system replicating a specific assembly, we now denote a system as a general *exponential replicator* if it replicates all patterns given some reasonable format that maps patterns to input assemblies. Let M denote a mapping from rectangular patterns over some alphabet ϕ to assemblies with the property that for any rectangular pattern w over ϕ, it must be that (1) $w = p(M(w))$ (The assembly representing pattern w must actually have pattern w), (2) all tiles in $M(w)$ with the same non-null label are the same active tile up to translation, and (3) the number of tiles in $M(w)$ is at most an additive constant larger than the size of w. Such a mapping is said to be a *valid format mapping* over ϕ. Constraint 2 is included so that a replicator's format mapping cannot encode additional information about a particular character, such as its location in the pattern, as all such characters would need to be represented by the same tile throughout the pattern. It may be reasonable to relax this constraint for some applications, but some limitation is important to model the idea of an input pattern in which each occurrence of a given character is locally indistinguishable from all other occurrences. We now define what constitutes an exponential pattern replicator system.

Definition 4 (*Exponential Replicator*) A system $T = (B, \tau)$ is an exponential pattern replicator for patterns over ϕ if there exists a valid format mapping M over ϕ such that for any rectangular pattern w over ϕ, $T = (B, \tau)$ exponentially replicates $M(w)$.

3 Replication of linear patterns

In this section, we focus on the replication of a linear assembly in two-dimensional space to ease the introduction of the extended mechanism presented in Sect. 4, which is the main result of this paper. We do not show in this section that the system satisfies the definitions of Sect. 2, as this is shown in Sect. 4. In this replication scheme, which occurs at temperature 1, some pre-formed linear patterned template assembly R is added to the replicating tile set T to compose the *initial seed batch* (Fig. 2). In general, the mechanism described follows the simple model outlined by von Kiedrowski et al. for template-directed self-replication. However, our scheme has a difference in that two types of products are formed: *terminal replicates* (*tr*) and *non-*

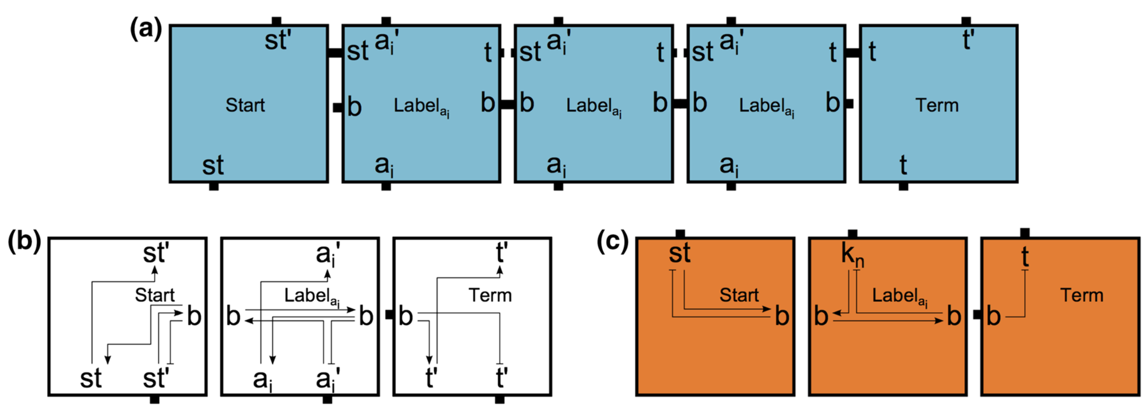

Fig. 2 The initial seed batch for replication of a 1 × n patterned template. $Label_{a_i}$ represents a tile labeled with element a_i in *alphabetA*. Glue a_i corresponds to this label and serves to identify the label of the tile to other tiles in the system. **a** General form of template to be replicated R **b** Tiles involved in formation of non-terminal replicates. **c** Tiles involved in formation of terminal replicates

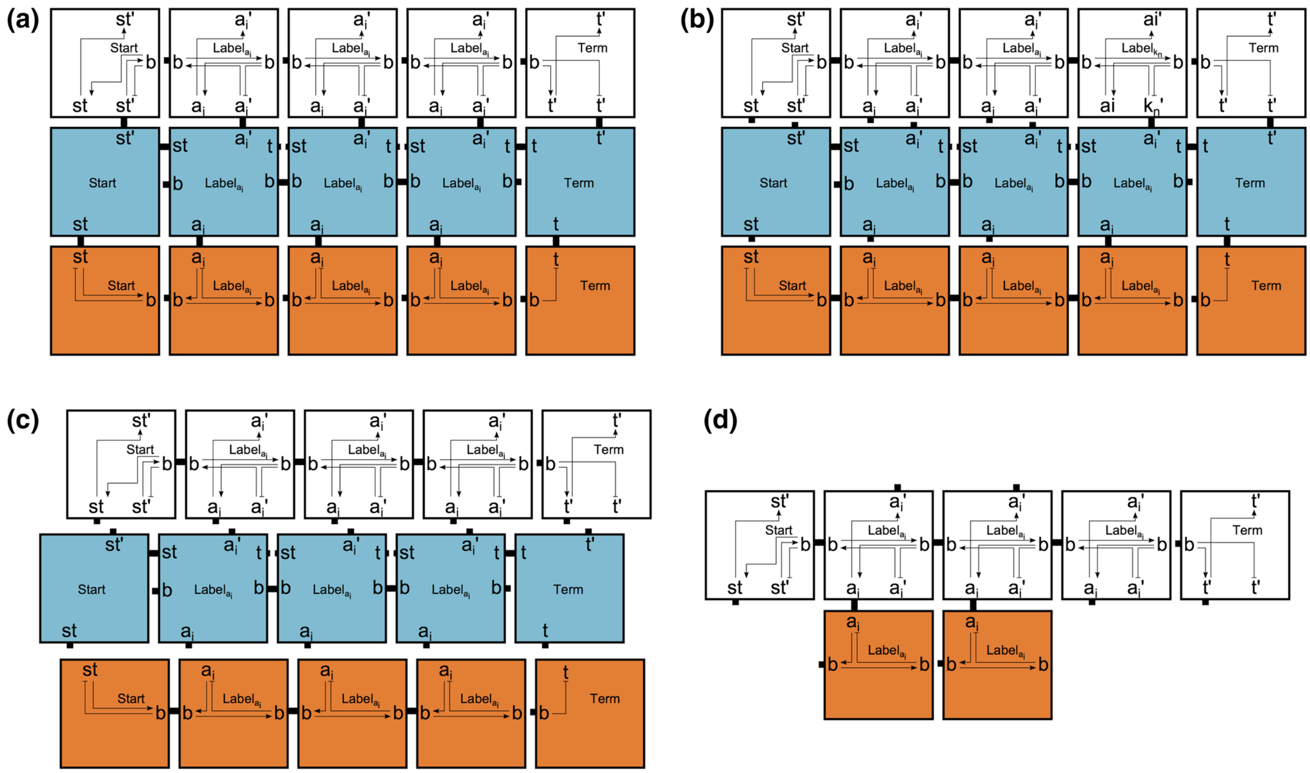

Fig. 3 **a** The *ntr* (*white*) and *tr* (*orange*) bind to the template assembly (*blue*) via label glues a_i, a_i', *st*, and *t*. **b** *b* glues are activated on the west face of the *ntr* and *tr Label* and *Term* tiles. The *b blue* on the *Start* tile is activated on the east side, which propagates a signal through the newly forming replicate to detach from the parent. **c** The newly-formed terminal replicate (*orange*) and non-terminal replicate (*white*) are completely detached from the parent template assembly. **d** North face label glues on an *ntr* template assembly are activated only upon binding of *tr* tiles to the south face of the template assembly. (Color figure online)

terminal replicates (*ntr*). While the pattern of each type of replicate is identical to that of the parent, each replicate type serves a different function. *Non-terminal replicates* may catalyze the formation of more product while *terminal* replicates serve as an inert final product and may not catalyze the formation of more product. Each *non-terminal replicate* may serve as a template for the formation of another *ntr* and a *tr* concurrently.

Upon addition of R to the replicating tile set, tiles involved in *non-terminal replicate* formation (white) may attach to the north face glues of the template: a_i', *st*, and *t*. Simultaneously, tiles involved in *terminal replicate* formation (orange) may attach to the south face glues of the template: a_i, *st*, and *t* (Fig. 3a). Upon binding, *b* glues are activated on the west face of each $Label_{a_i}$ and *Term ntr* or *tr* tiles. On *ntr* and *tr Start* tiles, *b* glues are activated on the east face upon binding to the template. Note that the template has no active signals. After binding of the *Start* tile, a signal is propagated from west to east along the newly forming replicate via the *b* glues. When a *Label* tile has bound the *b* glues on both its east and west faces, it may detach from the parent template (Fig. 3b). Following complete detachment of the replicates from the parent template (Fig. 3c), the *terminal replicate* is inert and may not undergo any further binding events. The *non-terminal replicate*, however, can continue to catalyze the formation of product. The *non-terminal replicate* has south face a_i glues exposed, allowing it to immediately serve as a template for the formation of a next-generation *terminal replicate* (Fig. 3d). Note that upon detachment, the north face glues a_i', *st*, and *t* of the *non-terminal replicate* are latent and are activated only upon the binding of *tr* tile to the south face of the *ntr*. This was designed so that *ntr* aggregates do not form, which would hinder the formation of *terminal replicates*. Following activation of the north face label glues, the *ntr* may serve as a template for the next-generation *ntr*s and *tr*s via the same mechanism as the original template assembly R.

Figure 4 summarizes the the process of tile attachment and dissociation and the importance of multiple signals in ensuring that complete dissociation does not occur until the copy is fully formed.

4 Replication of 2D patterns in two dimensions

We first informally discuss the mechanism for replication of 2D patterns in two dimensions with the tileset shown in

Fig. 6. The replication process described here can be summarized in three phases. In the first phase, *template disassembly*, a template R containing some pattern over some alphabet A is combined with the tile set that can replicate R. Initially, an inverted staircase cooperatively grows along the west face of R (Fig. 5, phase 1). The effect of this tile growth is that each row of the original assembly R has a unique number of tiles appended to its west side. These appendages are used in reassembly later in the replication process. As the inverted staircase structure grows, rows of the original template are signaled to detach from each other. In Phase 2, the detached rows of the input assembly are available to serve as templates for the formation of *non-terminal replicates* (Fig. 5, phase 2). As described in Sect. 3, two types of replicate products are formed: *terminal replicates* (*tr*) and *non-terminal replicates* (*ntr*). *Non-terminal replicates* may catalyze the formation of more product while *terminal* replicates serve as a final product and may not catalyze the formation of more product. After formation, this first generation of non-terminal replicates detach from the parent and enter phase 3. In phase 3, each *non-terminal replicate* may serve as a template for the formation of another *ntr* and a *tr* concurrently. The *tr* detaches from the parent upon completion and assembles, along with other terminal replicates, into a copy of R. Also during phase 3, when the new *non-terminal replicate* is fully formed, it may detach from the parent and begin producing replicates (Fig. 5, phase 3). Terminal replicates ensure that copies of R are produced, while non-terminal replicates ensure that copies of R are produced at an exponential rate.

We now present a more detailed synopsis of the replication mechanism. The 12 active tile types which comprise T are depicted in Fig. 6d–f. Note that the input pattern itself is not included in T. The input pattern to be replicated is of the form shown in Fig. 6c, and this, together with T, comprises the initial seed batch. The pattern is mapped onto this input via the composition of the *Label* signal tiles. Figure 6a shows the tile types for a binary alphabet, while Fig. 6b shows the tile type for some a_i of alphabet A which consists of elements $a_1, a_2, \ldots a_\phi$.

4.1 Template disassembly and first generation of replicates

Upon addition of the template assembly R to the replicating tile set T, an inverted staircase forms on the west side of R (Fig. 7a). Concurrently, an end cap attaches to the east side of R. Note that while the east-side end caps are attaching to R, it is possible that an *ntr* tile type (white) found in Fig. 6e may attach to the north side of an end cap, blocking the attachment of an endcap to a row. This does not adversely affect replication, because given a temperature of 2, the template will still disassemble and the end cap may attach to rows lacking end caps following this event. Also, given that the north face label glues a_i' of the northernmost template row are exposed, it is possible for this row to begin replicating immediately. In fact, this is necessary for the row immediately below the northernmost row to

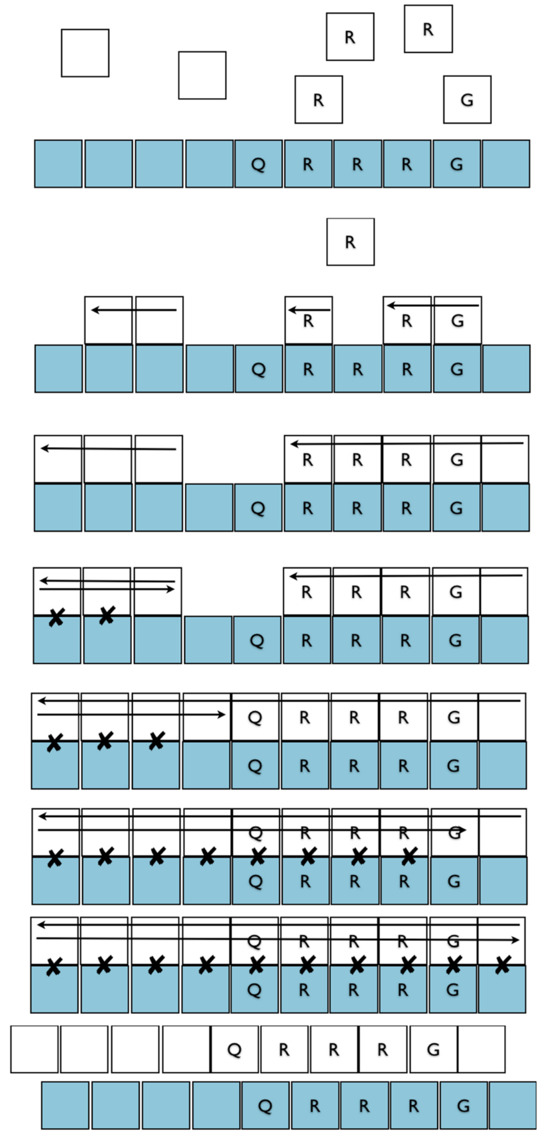

Fig. 4 From *top* to *bottom*: non-terminal replicate tiles (*white*) begin attaching to the template (*blue*). Upon attachment, each non-terminal replicate tile, with the exception of the westernmost tile, propagates a signal westward which waits if there is no tile to the west. Once a given signal has been propagated unbroken to the westernmost tile, a signal is propagated eastward. Tiles may only detach if both the westward and eastward signals have be propagated through the tile and onto the next. The eastward signal waits at a given tile, which does not detach, if there is no tile to the east. Signals are used in this way to ensure the copy does not dissociate before being fully formed. (Color figure online)

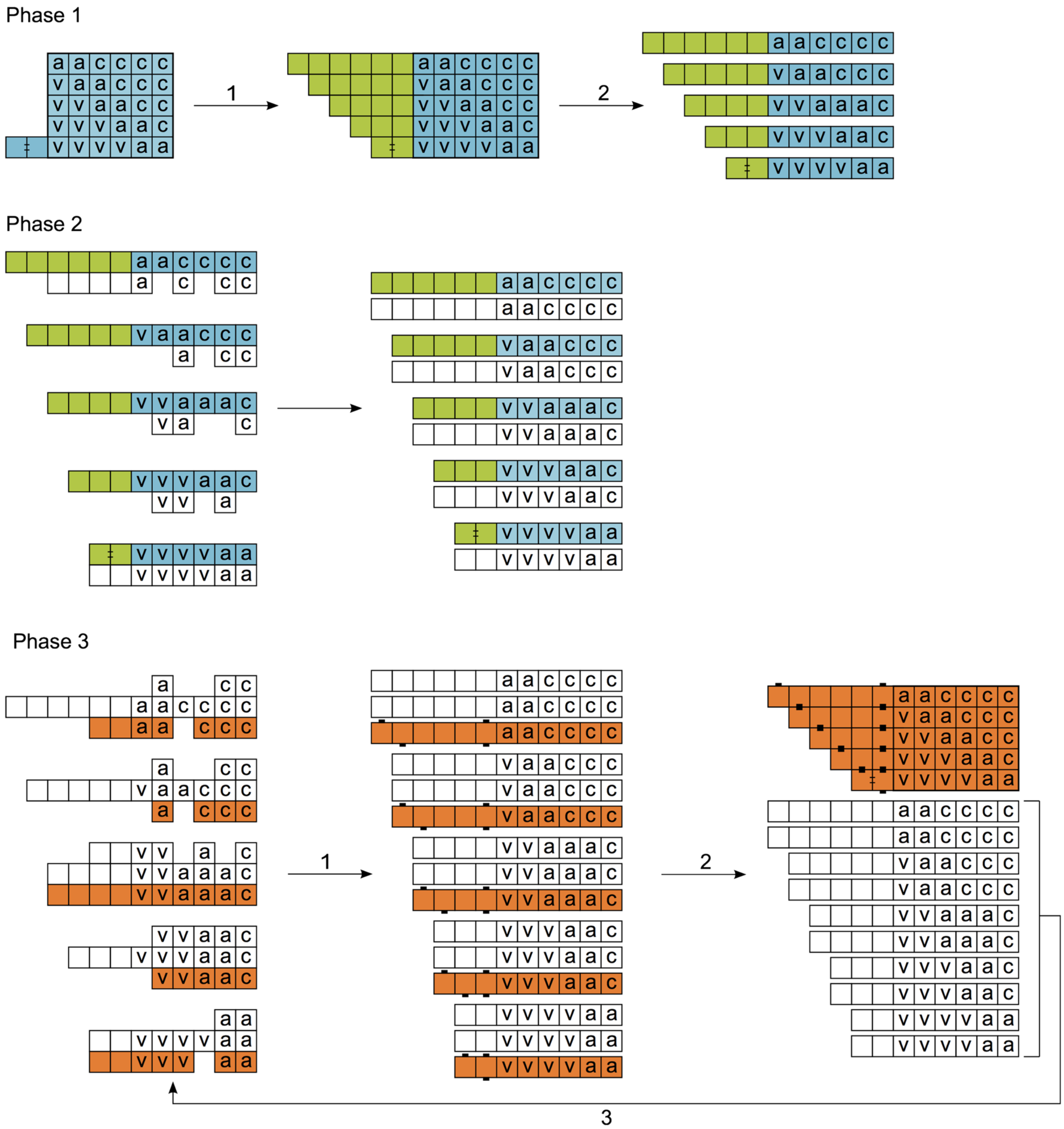

Fig. 5 The three phases shown above provide a general overview of the replication system described in this paper. In phase 1, an inverted staircase (*green*) cooperatively grows along the west face of the pattern to be replicated (*blue*). Upon completion of the staircase, the assembly splits into distinct rows. In phase 2, each of these distinct rows serves as a template for the production of a non-terminal replicates (*ntr*), shown in white, which has an identical pattern. In phase 3, these *ntr*s serve as templates for the production of identical *ntr*s and terminal replicates (*tr*), which are shown in *orange*. The *tr*s reassemble to form a copy of the original pattern while the *ntr*s continue to serve as templates for the production of more *tr*s and *ntr*s. (Color figure online)

detach. Any row s of R may release the row below it by turning off its south face glues (Fig. 7b). This can occur only if the row above s has activated the b glue on the westernmost tile of s. A signal is then propagated from west to east in row s via glue r and all south-face glues of s are turned off.

Following R disassembly, label glues a_i' are exposed on the north face of each row of the input assembly. Tiles involved in *ntr* formation (white) may attach along the north face of the template row (blue/green) (Fig. 8a). Following attachment, west face b glues are turned on. Once the westernmost *Label* tile has attached, appendage tiles may cooperatively attach, sending a signal via b glues from west to east and turning on r glues. (Fig. 8b). After the westernmost appendage tile has attached, a signal is propagated from west to east via glue r queueing label glues a_i' on the south face of the new *ntr* to turn *off*, thus detaching the *ntr* from its parent (Fig. 8c). Label glues a_i

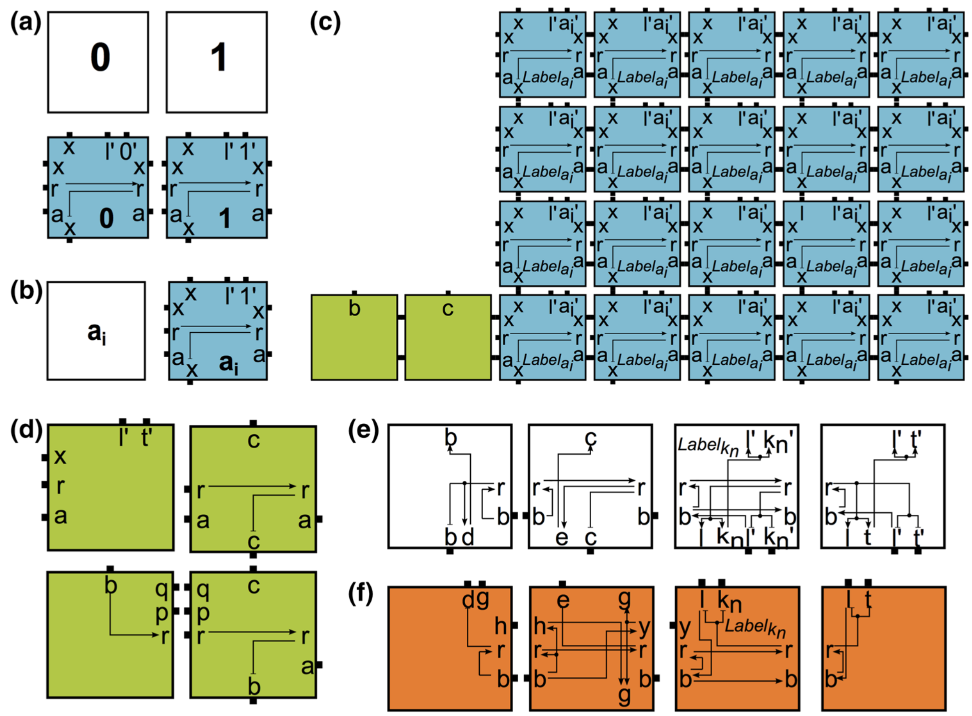

Fig. 6 **a** Input assemby tile types for a binary alphabet **b** The tile type for some a_i of alphabet A which consists of elements $a_1, a_2, \ldots a_\phi$. **c** General form of template to be replicated R **d** Tiles involved in inverted staircase construction and disassembly of the original template. **e** Tiles involved in formation of non-terminal replicates. **f** Tiles involved in formation of terminal replicates

are also queued *on*. These glues serve to generate a terminal replicate (*tr*) on the south face of the *ntr* (Fig. 8d). Following the detachment of the *ntr* and the parent template, the parent template is available to generate another *ntr*, while the first-generation *ntr* is immediately available to generate a *tr*.

4.2 Exponential replication and reassembly

After the formation of the first-generation *ntr*, replication is free to proceed exponentially. Glues on the south face of the *ntr* may bind label tiles from the *tr* tile set (Fig. 9a). Upon binding, *b* glues are turned *on* on the west face of the *tr* label tiles, allowing for the binding of appendage tiles on the western side of the growing *tr* assembly. Upon binding of the first appendage tile (Fig. 9b), a signal is propagated through the *tr* via glue *b* and *r* glues on the west faces of the *tr* tiles. After the next appendage tile binds, the *y* glue on the tile adjacent to it is activated, which activates two *g* glues on the north and south faces of the easternmost appendage tile (Fig. 9c). These *g* glues will assist in proper reassembly of each row into a correct copy of the template R. Also note that upon binding a *tr* tile, label glues a_i' on the north face of the *ntr* are turned *on*. This allows for synthesis of a new *ntr* on the north side of the parent *ntr* while a new *tr* is being formed on the south face. The synthesis of a new *ntr* from a parent *ntr* is not described in detail here, as it is very similar to the process described in Fig. 8. Upon attachment of the westernmost appendage tile, a north face *g* glue of the *tr* is turned *on* as well as a south face *g* glue on the tile immediately adjacent to it. Additionally, a signal is propagated from west to east along the *tr* via glue *r* and the north face glues of the *tr* are turned *off*. The *tr* then detaches from the parent *ntr* (Fig. 9d) and is available for reassembly into a copy of the original template R while the parent *ntr* is available to produce a new *ntr* on its north face and a new *tr* on its south face. The alignment of *g* glues enables the proper reassembly of the *terminal replicates* into a copy of R (Fig. 10).

Note that for the combination reactions shown in Fig. 10 to be feasible, we designated in Sect. 2 that at least one combinate must have multiplicity of at least c. This presents a bottleneck in the reassembly copies of R until the tr multiplicity is c. Therefore, shortly after R is introduced into the system, copies of R will form very slowly. Formation of copies of R will proceed exponentially once the population of trs has achieved sufficient multiplicity.

The detachment of the inverted staircase is not described here. If a signal cascade were designed such that upon the complete assembly of a copy of the original template pattern, the inverted staircase detached, it would be considered a waste product. The number of these waste assemblies grows proportionally to the number of replicates of R and the size of this waste product is proportional to the number of rows w of the input pattern. Similarly, if the replication process were somehow halted, and the copies of R harvested, the *ntr*s might also be considered waste. The number of these grows proportionally to the

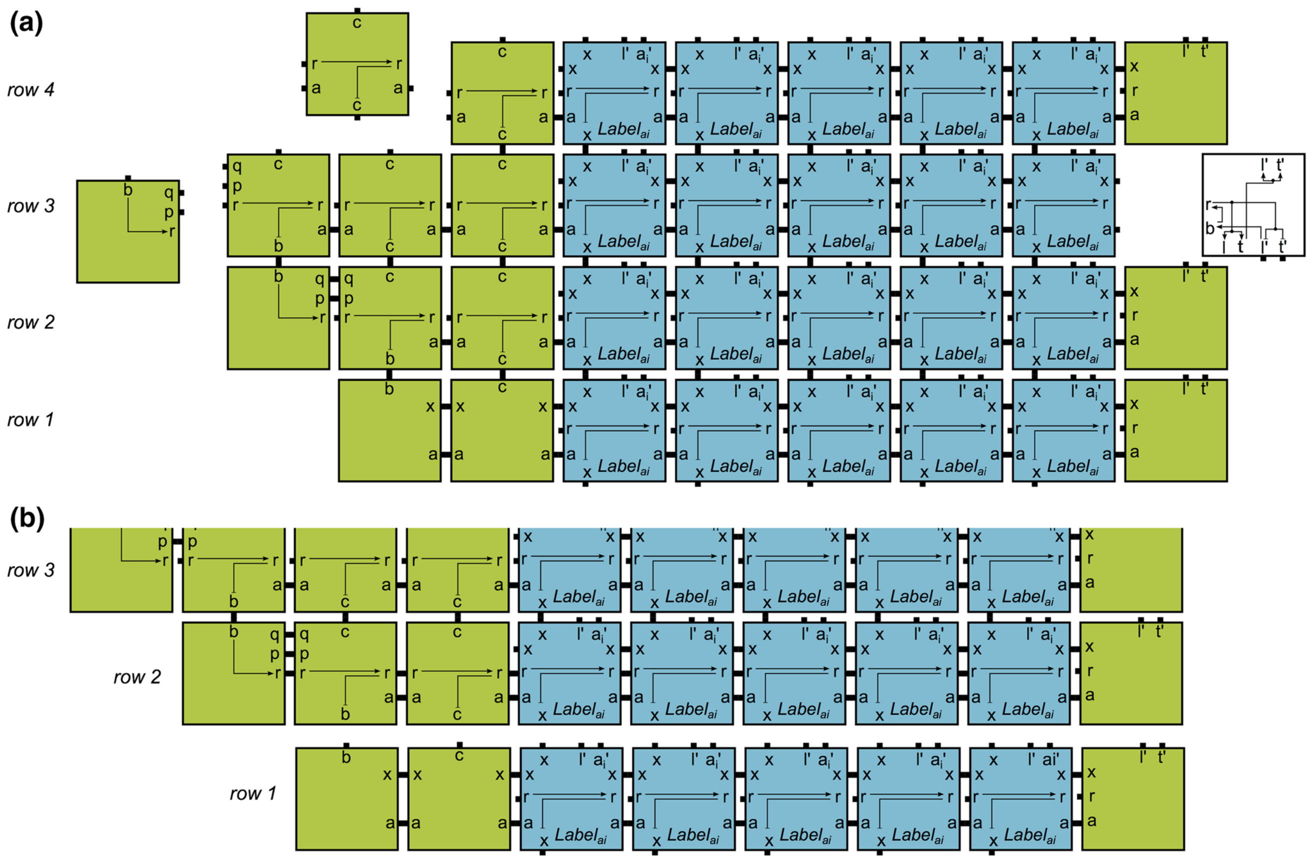

Fig. 7 **a** Growth of inverted staircase along the west face of R and a cap on the east face of R. **b** Row 1 is released after the b glue is activated on the westernmost tile of row 2

copies of R and to the number of rows in the template, w. Their size is proportional to the length ℓ and width w of the input pattern.

Theorem 1 For any alphabet A containing ϕ elements, there exists an exponential pattern replicator system $\Gamma = (T, 2)$ for patterns over A. Furthermore, the seed batch T consists of $O(\phi)$ distinct singleton active tile types with a total of $O(\phi)$ unique glues.

Proof To prove this we argue that the STAM system $(T, 2)$ defined by the tileset T depicted in Fig. 6b, c is an exponential pattern replicator. The valid format mapping M for the system is depicted in Fig. 6a.

We now argue that for any $w \times \ell$ pattern P, the assembly $A_P = M(P)$ derived by applying the format mapping described in Fig. 6a to pattern P is exponentially replicated by $(T, 2)$. First, by Lemma 5, the system satisfies the monotonicity requirement of the exponential replication. We therefore focus on the remaining requirement that for any positive integers n and c, the (τ, c)-transition distance from $T \bigcup A_P$ to some batch with at least n copies of an assembly with pattern P is $O(\log(n + c))$.

To show this, we construct a (τ, c)-transitional sequence of batches $\langle T \bigcup \{A_P\}, \ldots, B_{cleanBreak}, \ldots, B_{nReplicateRows}, \ldots, B_{nFinalRows}, \ldots, B_{nPatterns} \rangle$ with the property that batch $B_{cleanBreak}$ contains 1 non-terminal replicate of each row of the initial input assembly A_p, $B_{nReplicaterows}$ contains at least $n + c$ copies of the non-terminal replicate assembly for each row of the input assembly, $B_{nFinalRows}$ contains at least $n + c$ terminal replicates of each assembly row, and finally $B_{nPatterns}$ contains at least n copies of an assembly with pattern P. The construction for each segment of this sequence is depicted in Lemmas 1, 2, 3, and 4. From these Lemmas we get that the desired sequence of batches can be constructed with length at most $O(\log(n + c))$. $\square$

Lemma 1 For any $w \times \ell$ rectangular patterned input assembly P, let the batch $B_p = T \bigcup A_p$. For some sequence $\langle B_p, B_{p+1}, \ldots B_{cleanBreak} \rangle$, the (τ, c) transitional distance from B_p to $B_{cleanBreak}$ is $O(\ell + w\ell)$ where $B_{cleanBreak}$ contains at least one non-terminal replicate of each row within the input assembly.

Proof We first consider the cooperative binding of inverted staircase tiles. The entire inverted staircase forms

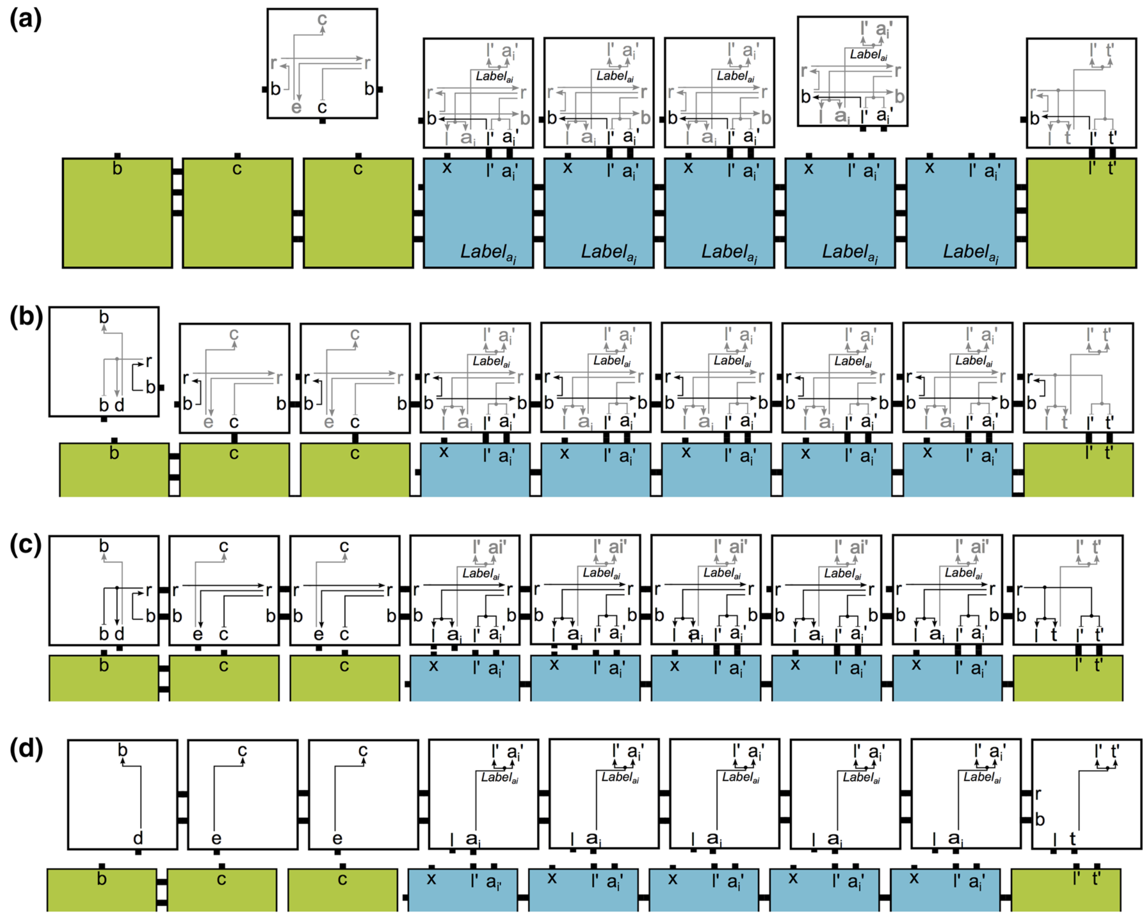

Fig. 8 The above sequence outlines details of the production of non-terminal replicates. For clarity, glues turned off and signals previously executed are not shown

cooperatively in $x = \frac{(\ell+1)(\ell+2)}{2} - 3$ combination reactions. Therefore, the (τ, c) transitional distance for the formation of the inverted staircase is x. For each row, signals must traverse west-to-east and then east-to-west across the entire row for the row beneath to completely detach. The number of glue-flip reactions required over the entire input assembly P for detachment of all rows from on another is $2((\ell-1)(w-1) + \frac{(\ell+1)(\ell+2)}{2} - 3)$. We then end up with input assembly rows with no other tiles attached. It follows from the analysis in case 1 that $O(w+\ell)$ transition steps are sufficient to generate clean non-terminal replicates.

Therefore, from any batch B_k, there exists a $O(w\ell + \ell^2)\,(\tau, c)$ transitional distance to batch $B_{cleanBreak}$, where batch $B_{cleanBreak}$ contains a clean non-terminal replicate of each row of the input assembly . □

Lemma 2 Consider that batch B_k contains one clean non-terminal replicate of each input assembly row. Then there exists a (τ, c) transitional distance of $O((\ell + w)\log(n + c))$ for the batch transition sequence $\langle B_k, B_{k+1}, \ldots B_{ReplicateRows}\rangle$ such that $B_{ReplicateRows}$ contains $n + c$ copies of non-terminal replicates of each input assembly row.

Proof For any newly-assembled *ntr* to generate an identical offspring *ntr*, it must first bind *terminal replicate* tiles in order to activate its north-face glues which serve as templates to bind *ntr* tiles. This is a one-time activation event, after which an *ntr* may generate unbounded copies of identical *ntrs*. For any newly generated *ntr*, at most $w + \ell + 2$ combination reactions and $2w + \ell + 4$ glue-flip reactions must occur to activate the north-face glues on the parent *ntr*. Once an *ntr* has been activated, *ntr* label tiles and the easternmost cap tile bind in $w + 1$ combination reactions. Following these combination reactions, a series of at most $8w + 7\ell + 4$ glue-flip and combination reactions are required to fully connect the newly-formed *ntr* and detach it from the parent input assembly row. In total, there is a $12w + 9\ell + 11$, or $O(\ell + w)\,(\tau, c)$ transitional distance *ntr* exist for each *ntr* in a batch to be activated and generate an identical *ntr*. Thus, after every $O(\ell + w)$ transitions, the population of *ntrs* doubles. Therefore, there is a $O((\ell + w)\log(n + c))\,(\tau, c)$

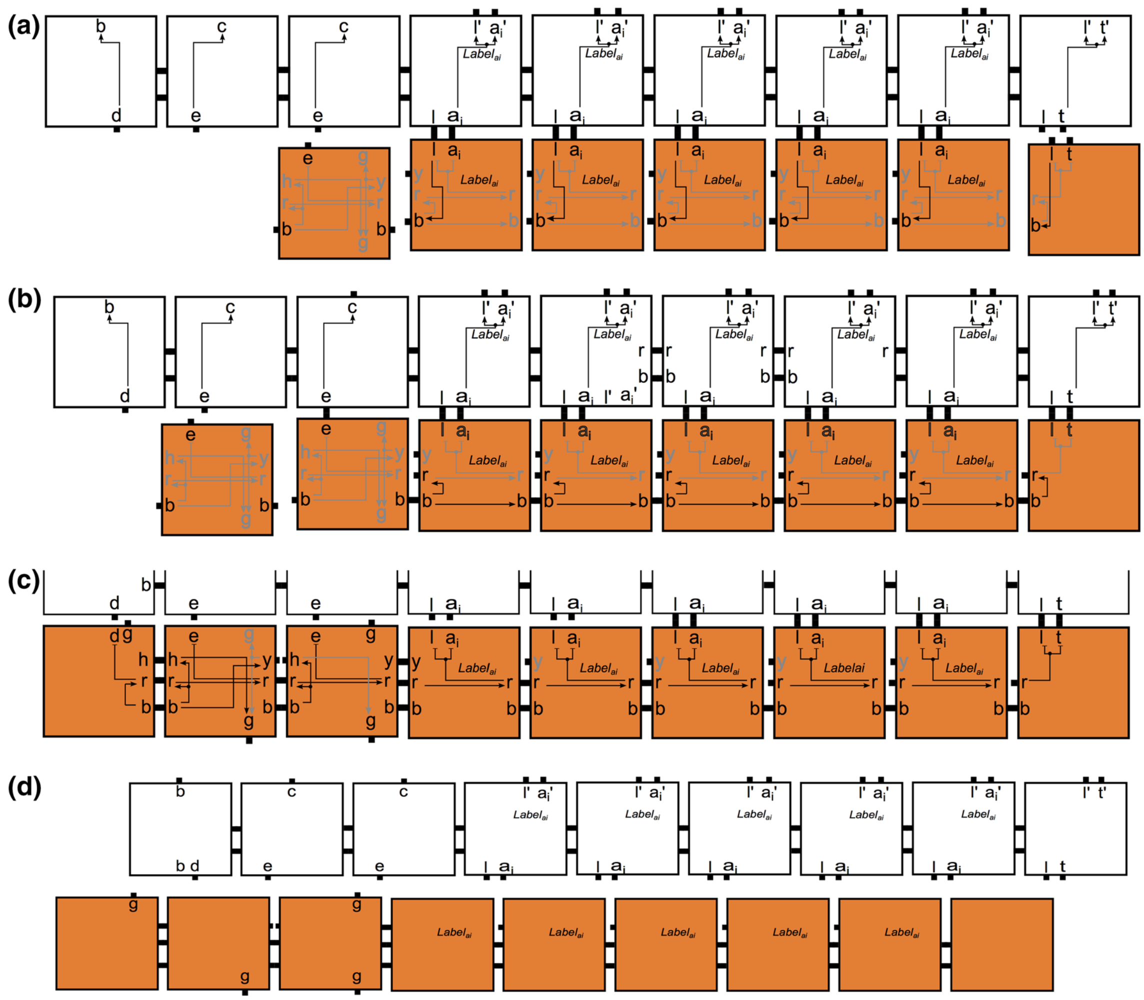

Fig. 9 The above sequence shows details of the formation of terminal replicates. For clarity, glues turned off and signals executed during template disassembly are not shown

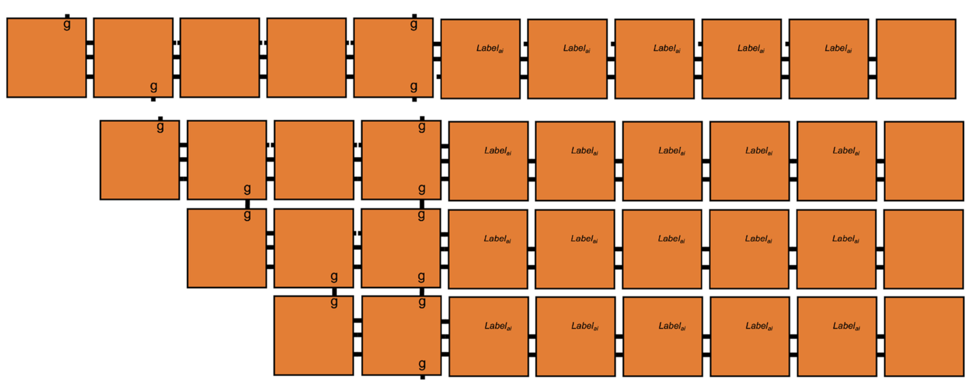

Fig. 10 Terminal replicates reassemble into a copy of R

transitional distance to achieve $n+c$ copies of a non-terminal replicates of each input assembly row. $\square$

Lemma 3 Consider that batch B_k has n clean non-terminal replicates of each row of the input assembly. There exists a batch sequence $\langle B_k, B_{k+1}, \ldots B_r \rangle$ with a (τ, c) transitional distance of $O(w+\ell)$ from B_k to $B_{FinalRows}$ where batch $B_{FinalRows}$ contains $n+c$ terminal replicates of each row.

Proof Terminal replicate label tiles, the easternmost capping tile, and the tag tiles must bind with a non-terminal assembly in at most $w + \ell + 2$ combination reactions. Following these tile attachments, $8w + 7\ell + 5$ glue-flip and combination reactions are required to fully connect the newly-formed *tr* and detach it from the parent *ntr*. Therefore, $9w + 8\ell + 7$ parallel reactions exist to produce n *terminal replicates* from n *non-terminal replicates*. □

Lemma 4 Consider that batch B_k has $n + c$ terminal replicates of each row. There exists a batch sequence $\langle B_k, B_{k+1}, \ldots B_{nPatterns}\rangle$ with a (τ, c) transitional distance of $O(\ell)$ where batch $B_{nPatterns}$ contains n identical assemblies a where $P(a)$ is identical to the input assembly.

Proof Upon detachment from the parent, terminal replicates have g glues oriented such that they may attach to the correct neighbors within the patterned assembly. For these terminal replicates to combine into assembly that has an identical pattern to the input assembly, ℓ rows must attach to one another in $\ell - 1$ combination reactions. Therefore, $O(\ell)$ parallel batch transitions exist for batch B_k to transition to batch $B_{nPatterns}$. □

Lemma 5 Consider the seed batch B_p where $B_p = T \bigcup A_p$. B_p grows monotonically toward G where G contains at least n copies of an assembly with pattern P

Proof For any valid batch sequence $\langle B_p, B_1, \ldots, B_k\rangle$, let the (τ, c) transitional distance from batch B_k to B_r be x where B_r contains n identical assemblies a where $P(a)$ is identical to the initial input pattern. For some batch sequence $\langle B_k, B_{k+1}, \ldots, B_\ell\rangle$, let the (τ, c) transition distance from B_ℓ to B_r be y. We transition B_k and B_ℓ to the nearest 'perfect' batches P_k and P_ℓ, respectively, where the ntrs that comprise P_k and P_ℓ have no tiles attached to their north or south faces. This means that during the transition from B_k to P_k or from B_ℓ to P_ℓ, any partially formed trs or ntrs on the ntr templates are completed and detach. Therefore, batches P_k and P_ℓ will be comprised of some number of completed trs (some may be combined with one another to form copies of a or partial copies of a) and ntrs with no other tiles attached. Of these ntrs, there are two classes: passive ntrs and active ntrs. Passive ntrs have no more active signals and may serve as a template for the formation of a tr and an ntr concurrently, which form independently of each other. Active ntrs are those that have just been released from their parent ntr and have not yet served as a template. These must first serve as a template for a tr. Upon formation of this first child tr, the active ntr will become passive. Because, an ntr must first serve as a template for at least one tr before an ntr, for any perfect batch P_i, the number of trs in the batch must be at least as great as the number of passive ntrs. The (τ, c) transitional distance from $\langle B_i, B_{i+1}, \ldots, P_i\rangle$ is $O(\ell + w)$.

Because B_k can transition into B_ℓ, the numbers of passive ntrs, active ntrs, and trs must each be at least as large in P_ℓ as in P_k when $P_k \neq P_\ell$. Therefore, when $P_k \neq P_\ell$, $x \leq y$, thus satisfying the monotonic growth requirement.

When $P_k = P_\ell$, we consider the transition $\langle B_k, B_{k+1}, \ldots, P_k, \ldots, Pk + 1\rangle$. Because B_k may transition into B_ℓ, B_ℓ can mimic the path of B_k from $P_k = P_\ell$ to $Pk + 1$ in at most as many steps as B_k. □

5 Future work

The results of this paper provide several directions for future work. One interesting problem is the replication of shapes in the STAM, or more specifically, patterned shapes. One might imagine a mechanism similar to the one presented in this paper but where the growth of the inverted staircase is preceded by a "rectangularization" of the shape to be replicated. The replication of a cuboid is conceivable by extending the mechanism of template disassembly and reassembly presented in Sect. 4 to three dimensions where layers of the cuboid might be separated, replicated, and the replicates reassembled. Precise replication of a certain number of copies could also be possible, as was considered in Abel et al. (2010).

Another direction for future work is studying the extent to which staged self-assembly systems can be simulated by non-staged active self-assembly systems such as the signal tile model. In Demaine et al. (2008) efficient staged algorithms are developed to assemble linear structures, while a signal tile system achieves a similar result in Padilla et al. (2013). Shape replication through stages and RNA based tiles are used to replicate general shapes in Abel et al. (2010), while this paper and future work suggests similar results may be obtained with signal tiles. Can the complexity of the mixing algorithm of a staged assembly algorithm be encoded into a signal tile system of similar complexity? As a first step towards such a simulation we might consider the case of 1D assemblies. Can the efficient construction of labeled linear assemblies through staging shown in Demaine et al. (2011) be efficiently simulated with a signal tile system?

A final direction for future work involves the simulation of the signal tile model through a passive model of self-assembly such as the abstract or two-handed tile assembly model (Cannon et al. 2012). Recent work has shown how restricted classes of signal tile systems can be simulated by passive 3D systems (Hendricks et al. 2013). The ability for signal tile systems to perform *fuel-efficient* computation was shown to be achievable within passive 2D tile assembly given the added power of negative force glues

(Schweller and Sherman 2013). Is it possible to simulate *any* signal tile system with the use of negative glues? Can this be done in 2D?

Acknowledgments We would like to thank Jennifer Padilla for helpful discussions regarding the current state of experimental DNA implementations of signal tile systems. This research was supported by National Science Foundation Grant CCF-1117672.

References

Abel Z, Benbernou N, Damian M, Demaine E, Demaine M, Flatland R, Kominers S, Schweller R (2010) Shape replication through self-assembly and RNase enzymes. In SODA 2010: Proceedings of the 21th snnual ACM-SIAM symposium on discrete algorithms, Austin, Texas. Society for industrial and applied mathematics

Cannon S, Demaine ED, Demaine ML, Eisenstat S, Patitz MJ, Schweller R, Summers SM, Winslow A (2012) Two hands are better than one (up to constant factors). Arxiv, preprint arXiv:1201.1650

Demaine ED, Demaine ML, Fekete SP, Ishaque M, Rafalin E, Schweller RT, Souvaine DL (2008) Staged self-assembly: nanomanufacture of arbitrary shapes with $O(1)$ glues. Nat Comput 7(3):347–370

Demaine ED, Eisenstat S, Ishaque M, Winslow A (2011) One-dimensional staged self-assembly. In Proceedings of the 17th international conference on DNA computing and molecular programming, DNA'11, pp 100–114

Hendricks J, Padilla JE, Patitz MJ, Rogers TA (2013) Signal transmission across tile assemblies: 3d static tiles simulate active self-assembly by 2d signal-passing tiles. Proceedings of the 19th international conference on DNA computing and molecular programming (DNA 19), vol 8141. Tempe, Arizona, pp 90–104

Kao M.-Y, Schweller RT (2006) Reducing tile complexity for self-assembly through temperature programming. In SODA 2006: Proceedings of the 17th annual ACM-SIAM symposium on discrete algorithms, pp 571–580

Keenan A, Schweller R, Zhong X (2013) Exponential replication of patterns in the signal tile assembly model. Proceedings of the 19th international conference on DNA computing and molecular programming (DNA 19), vol 8141. Tempe, Arizona, pp 118–132

Lincoln T, Joyce G (2009) Self-sustained replication of an RNA enzyme. Science 323(5918):1229–1232

Marchal P (1998) John von neumann: the founding father of artificial life. Artif Life 4(3):229–235

Padilla J, Liu W, Seeman N (2011) Hierarchical self assembly of patterns from the Robinson tilings: DNA tile design in an enhanced tile assembly model. Nat Comput 11(2):323–338

Padilla JE, Patitz MJ, Pena R, Schweller RT, Seeman NC, Sheline R, Summers SM, Zhong X (2013) Asynchronous signal passing for tile self-assembly: fuel efficient computation and efficient assembly of shapes. Proceedings of the 2013 international conference on unconventional computation and natural computation

Patzke V, von Kiedrowski G (2007) Self-replicating sytems. ARKIVOC 5:293–310

Paul N, Joyce GF (2002) A self-replicating ligase ribozyme. PNAS 99(120):12733–12740

Reif JH, Sahu S, Yin P (2005) Complexity of graph self-assembly in accretive systems and self-destructible systems. DNA computing., LNCSSpringer, Heidelberg, pp 257–274

Schulman R, Yurke B, Winfree E (2012) Robust self-replication of combinatorial information via crystal growth and scission. PNAS 109(17):6405–6410

Schweller R, Sherman M (2013) Fuel efficient computation in passive self-assembly. Proceedings of the 24th annual ACM-SIAM symposium on discrete algorithms (SODA 2013). Louisiana, New Orleans, pp 1513–1525

Tjivikua T, Ballester P, Rebek J Jr (1990) Self-repllicating system. J Am Chem Soc 112(3):1249–1250

von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. Angew Chemie (International Edition in English) 25(10):932–935

Winfree E (1998) Algorithmic self-assembly of DNA. PhD Thesis, California Institute of Technology

Zielinski W, Orgel L (1987) Autocatalytic synthesis of a tetranucleotide analogue. Nature 327:346–347