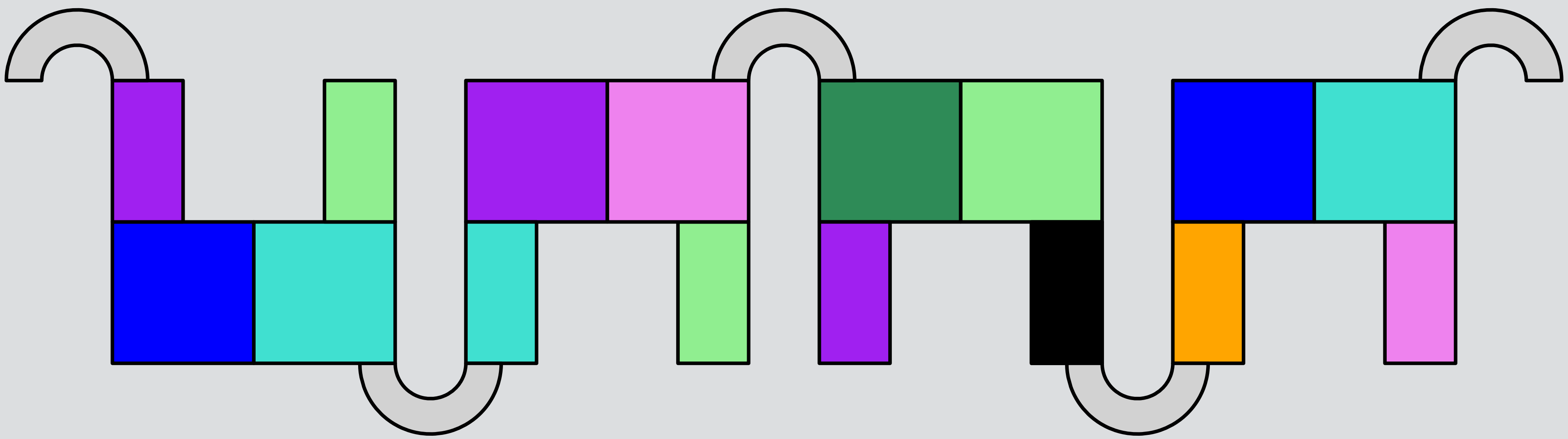
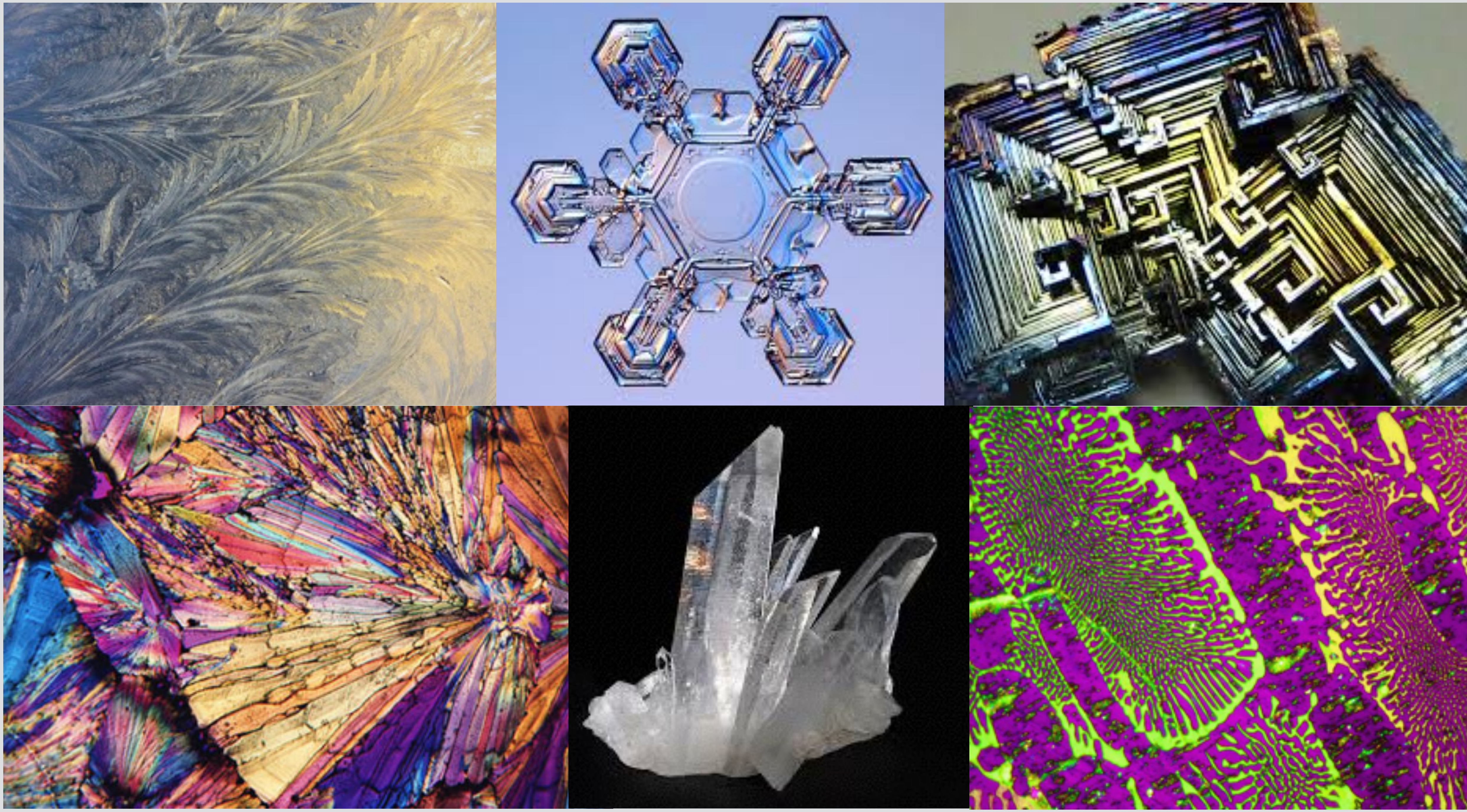
The Limits of a Simple Model of Active Self-Assembly



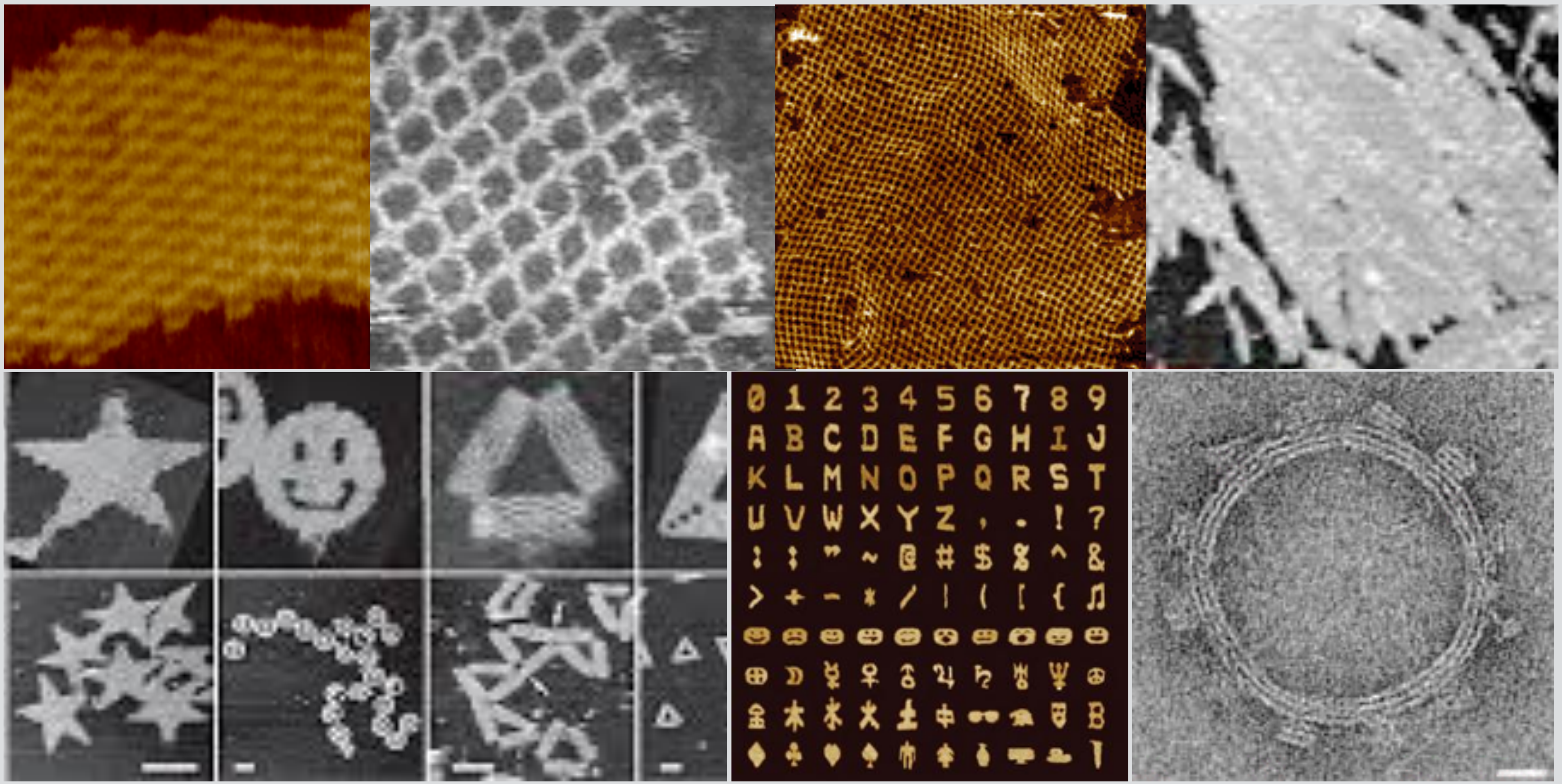
Andrew Winslow

UNIVERSITÉ LIBRE DE BRUXELLES

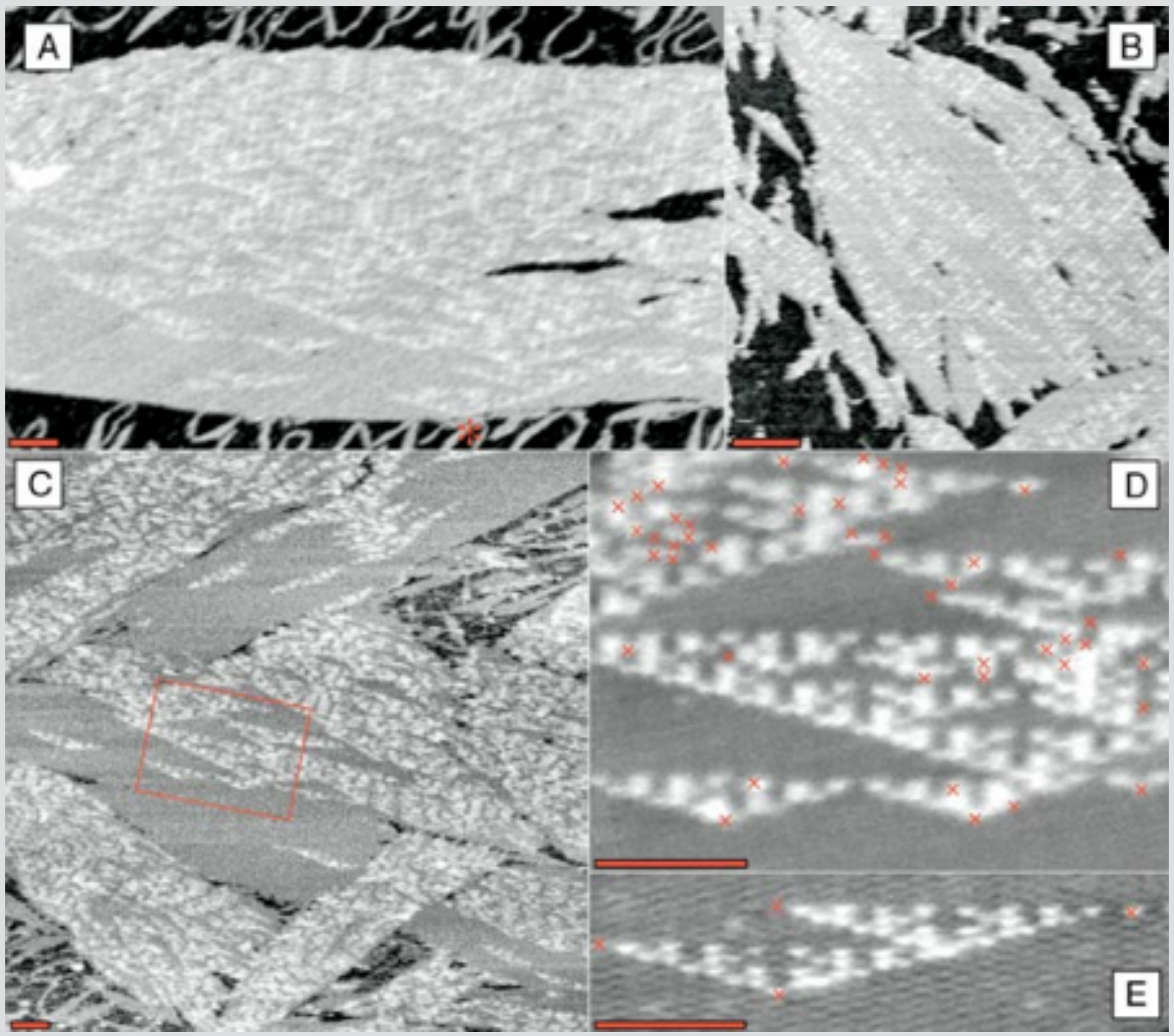
Natural Nanoscale Self-Assembly



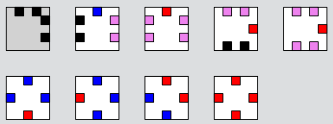
Synthetic Nanoscale Self-Assembly



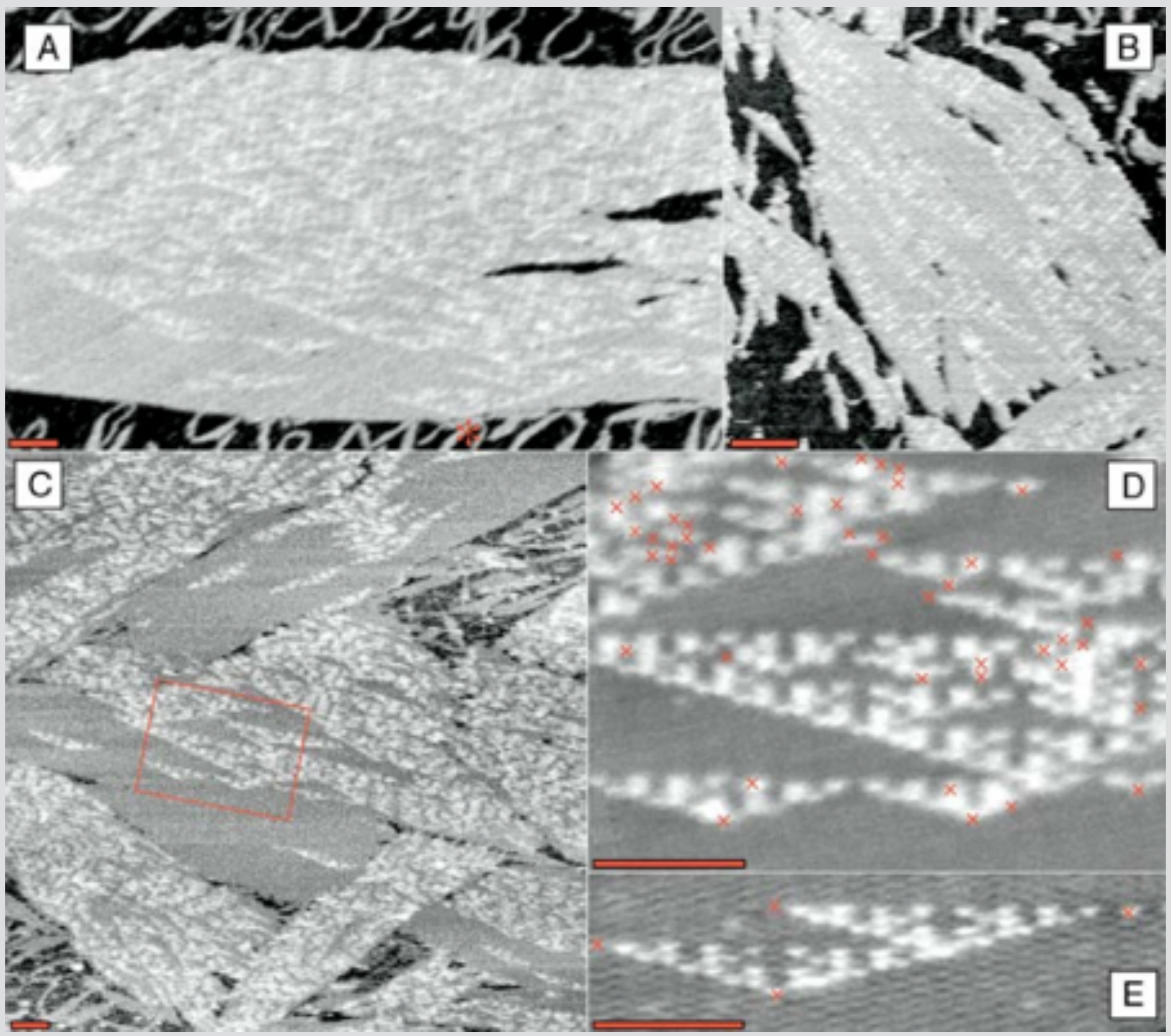
Implementation (in DNA)



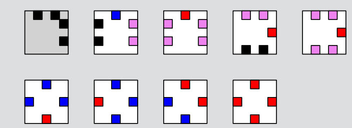
Theoretical model (aTAM)

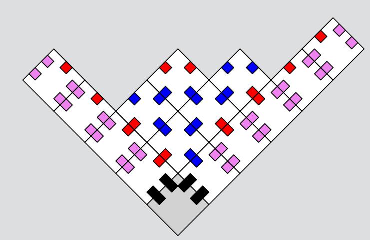


Implementation (in DNA)

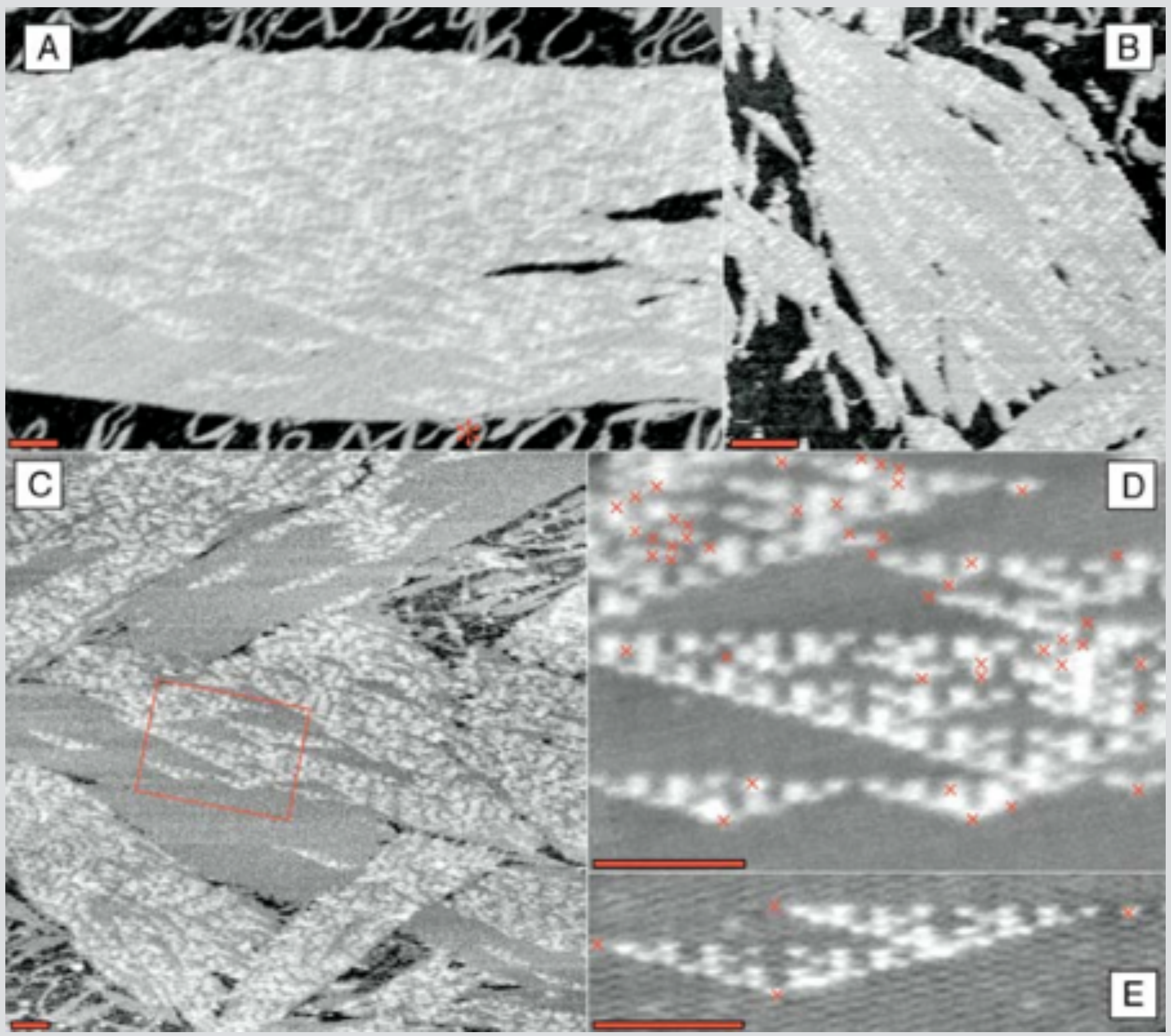


Theoretical model (aTAM)

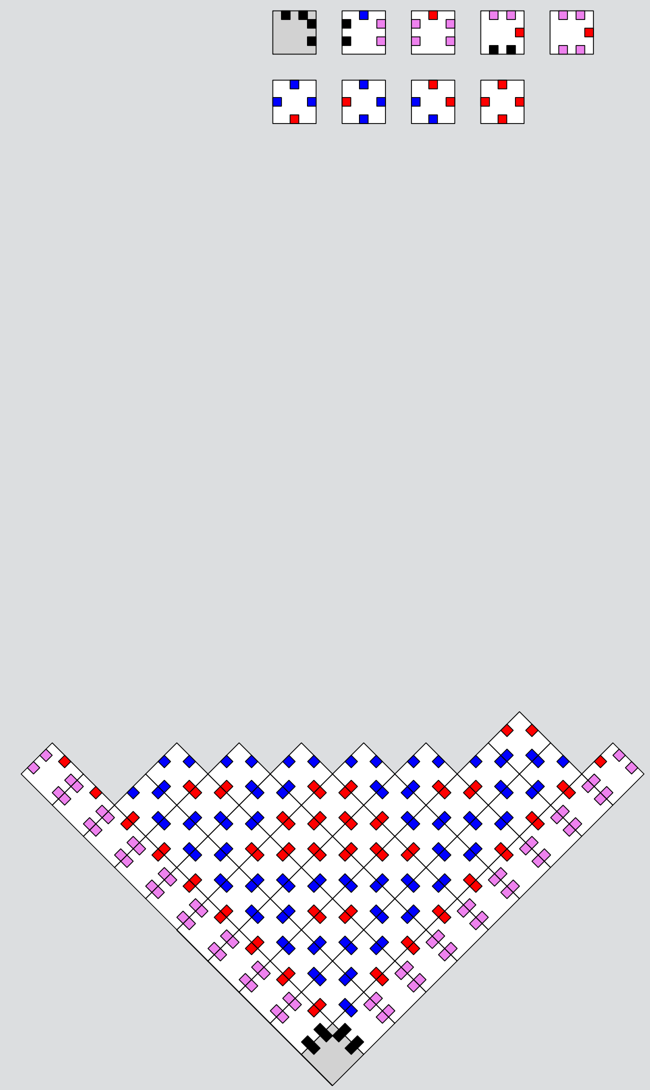




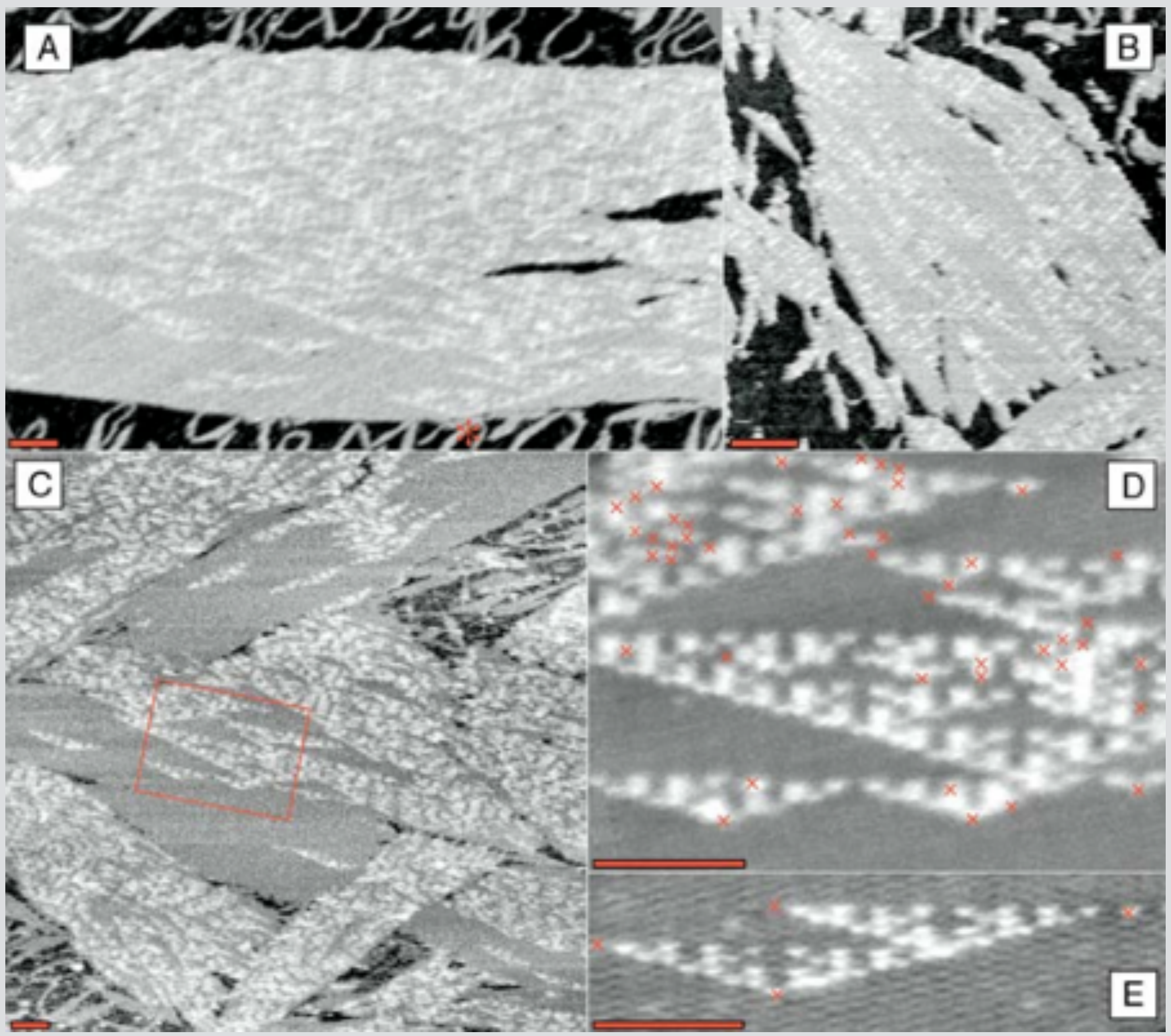
Implementation (in DNA)



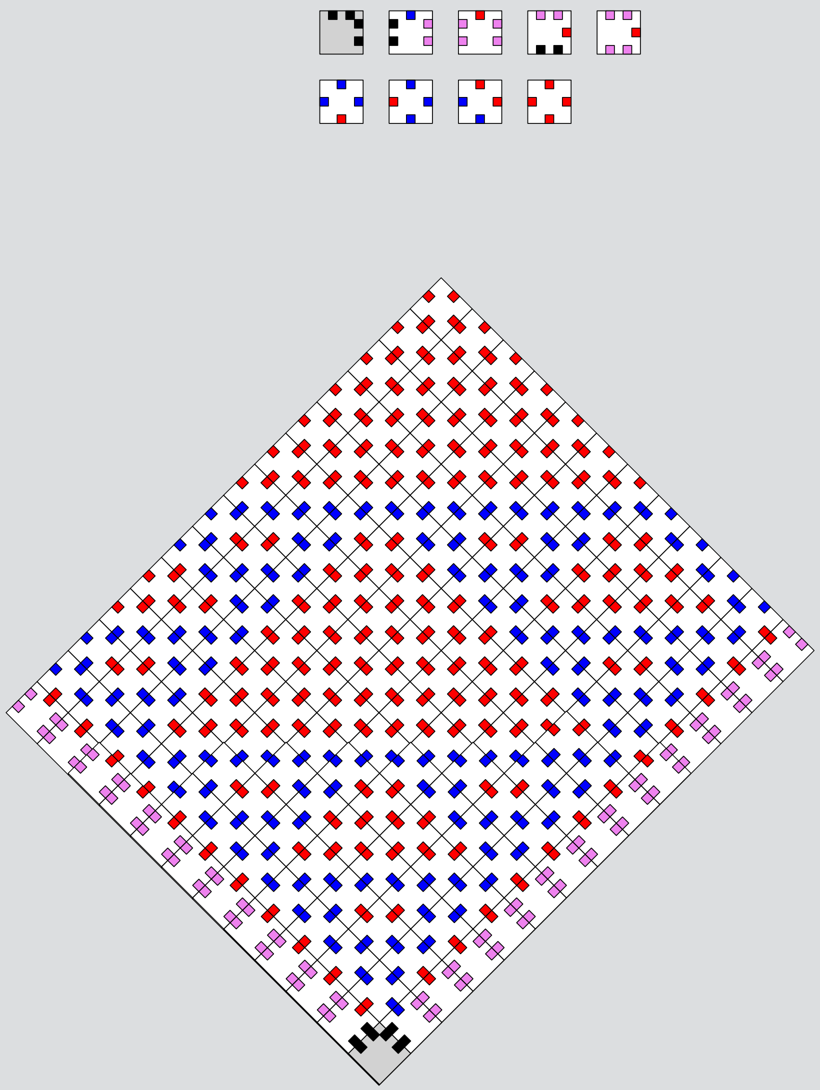
Theoretical model (aTAM)



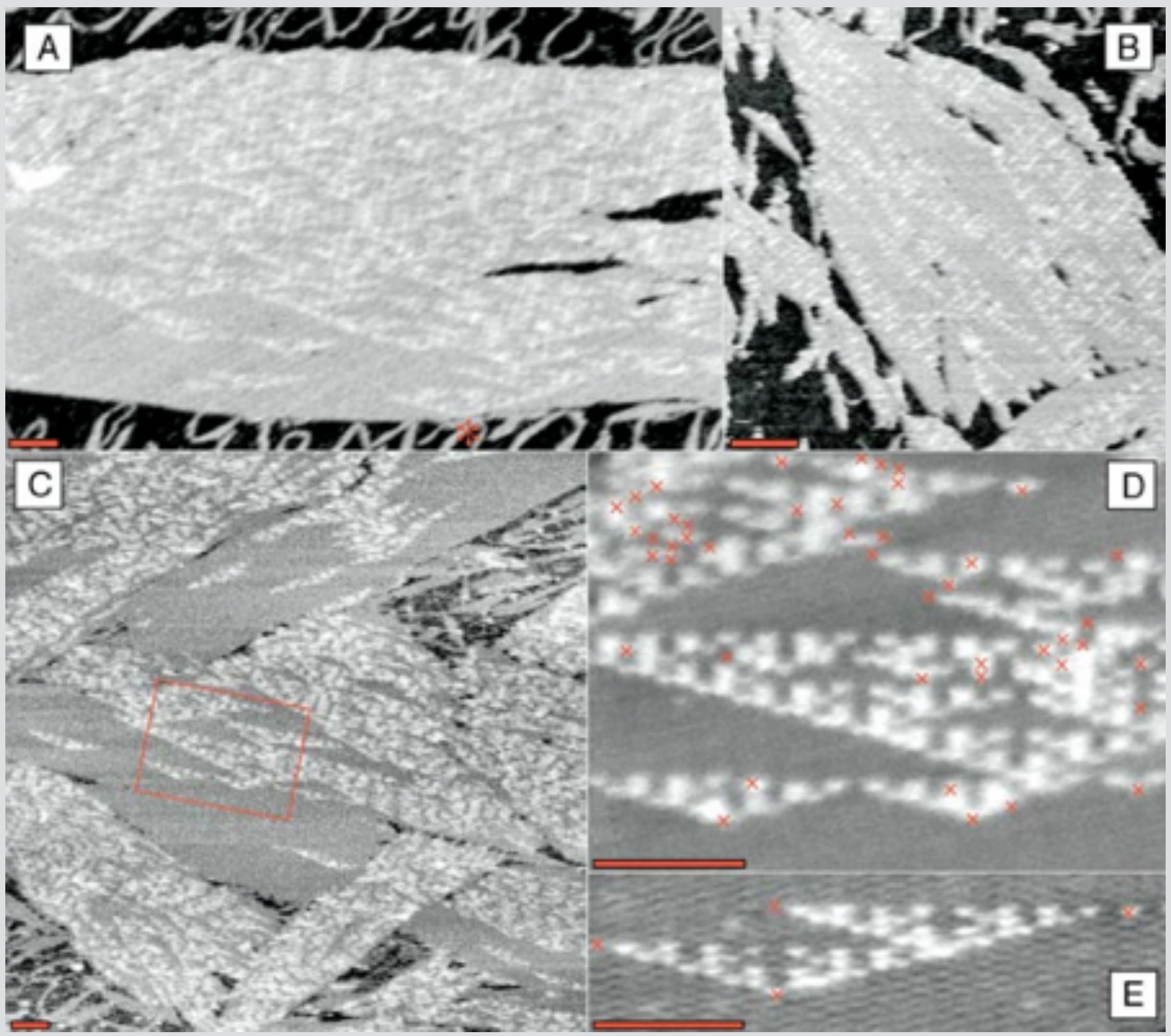
Implementation (in DNA)



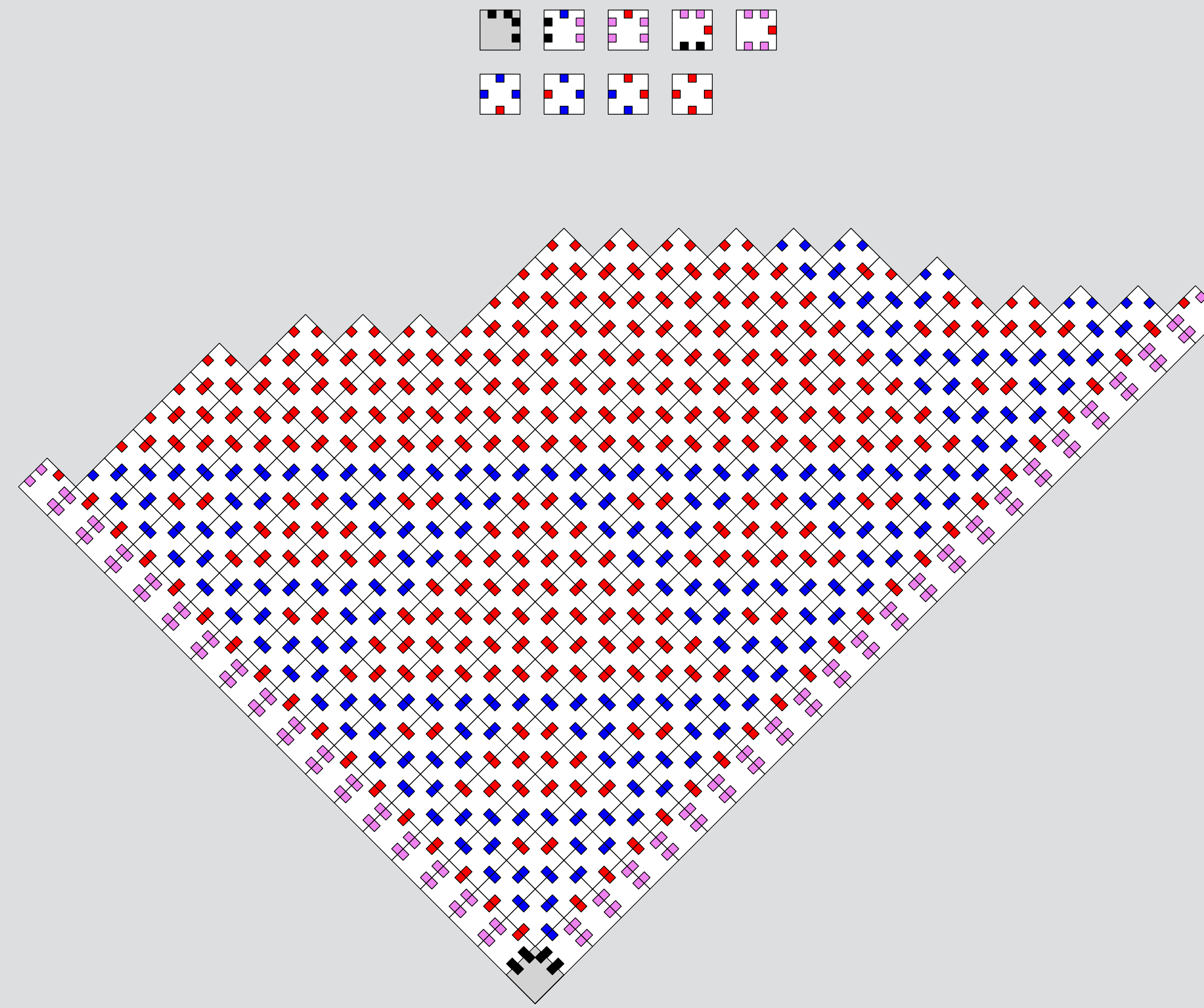
Theoretical model (aTAM)



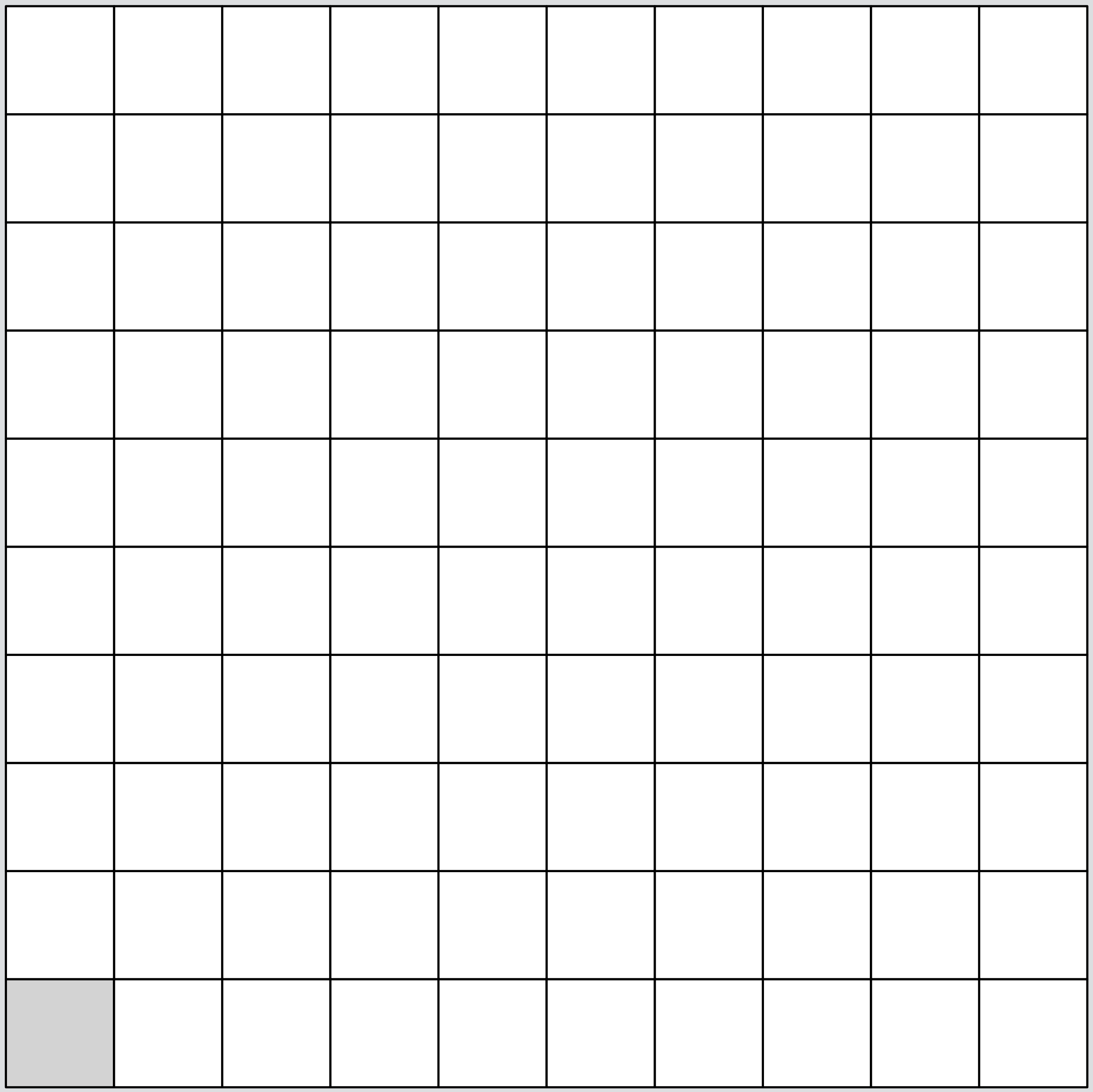
Implementation (in DNA)



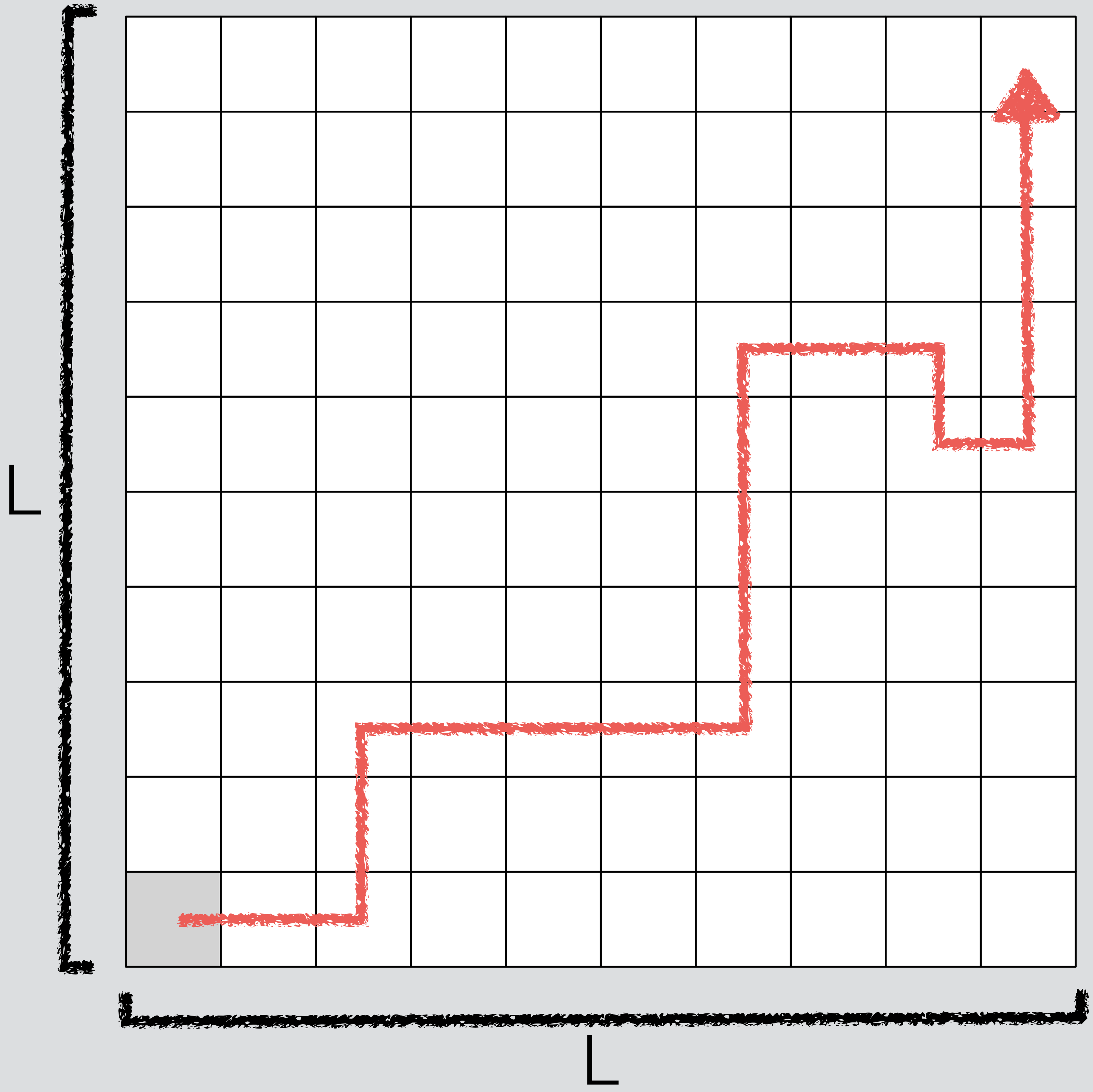
Theoretical model (aTAM)



A growth rate limitation

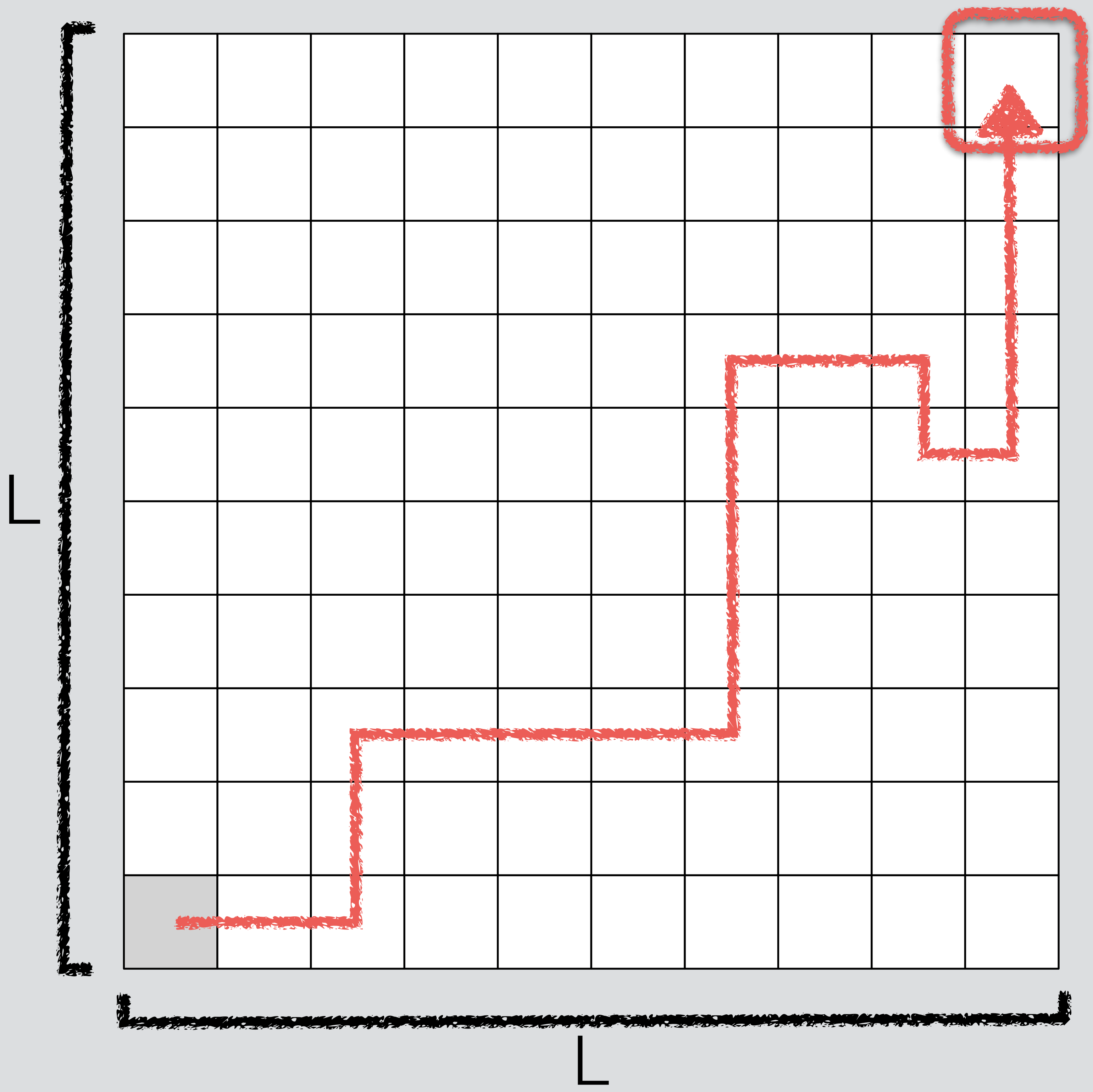


A growth rate limitation



path length ≥ L

A growth rate limitation



path length ≥ L

at least L tile prior tile placements

Growth rates

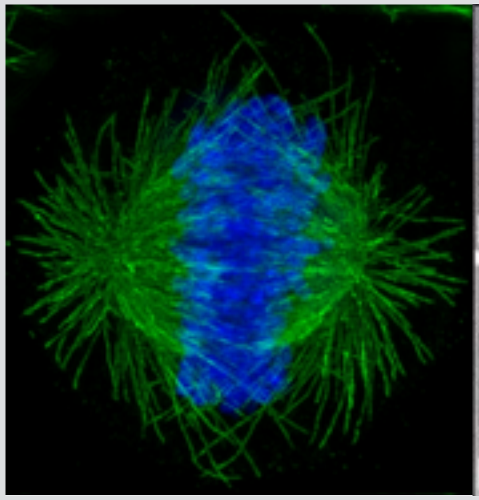
- Takes Ω(diameter) = $\Omega(n^{0.5})$ expected time for n particles to assemble on a square lattice.

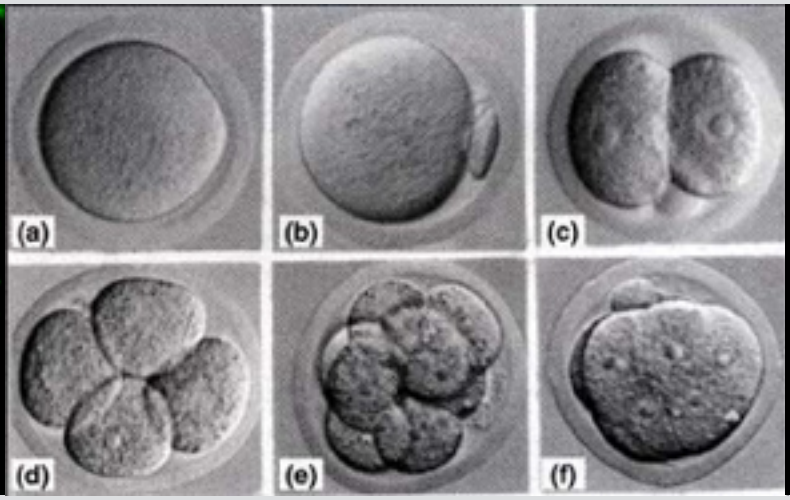
- So growth rate is $O(n^2)$.

- Similar polynomial bounds for other lattices.

Growth rates

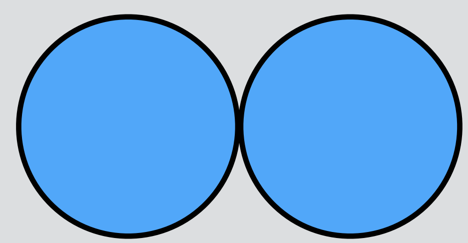
- Takes $\Omega(\text{diameter}) = \Omega(n^{0.5})$ expected time for n particles to assemble on a square lattice.
- So growth rate is $O(n^2)$.
- Similar polynomial bounds for other lattices.
- But not all nanoscale growth is polynomial!

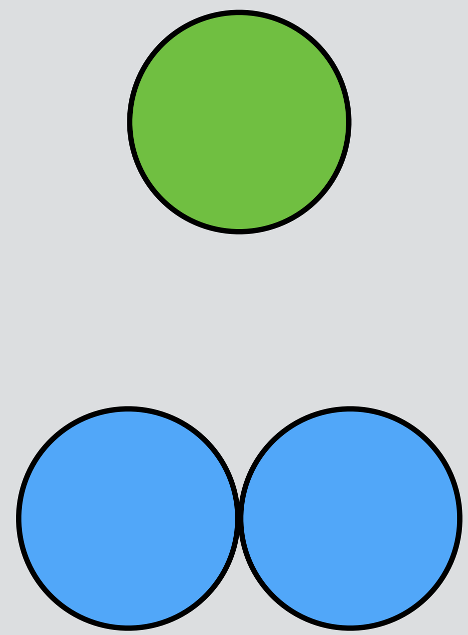


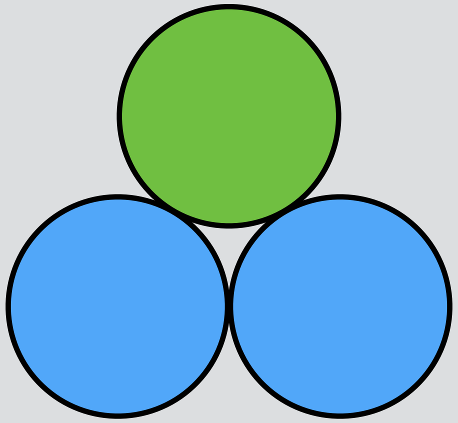


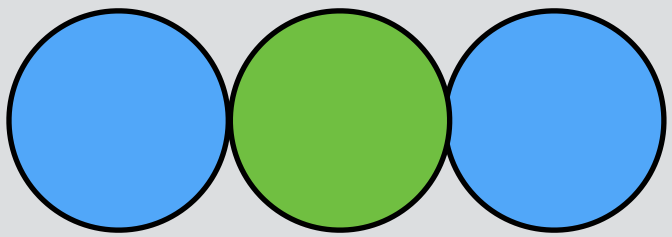
Passive vs. Active Self-Assembly

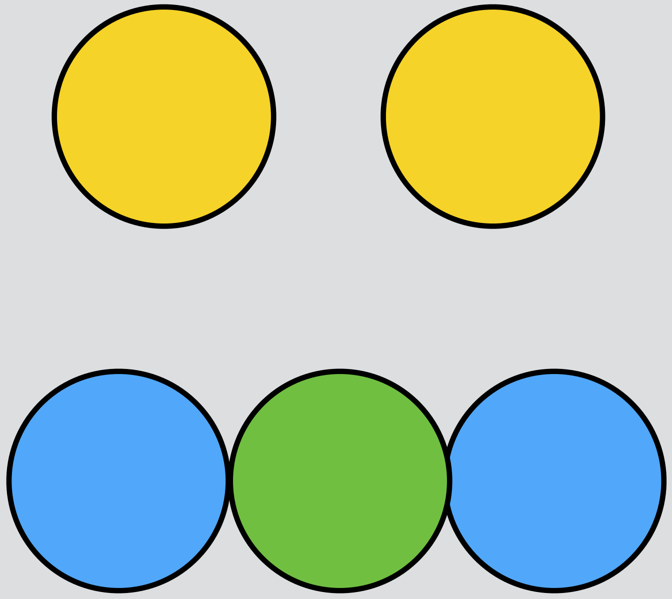
- Most current DNA self-assembly uses crystal-like **passive** growth: bonds and geometry do not change.
- Some natural systems use **active** growth: bonds and geometry change.
- Active growth enables exponential growth rates.

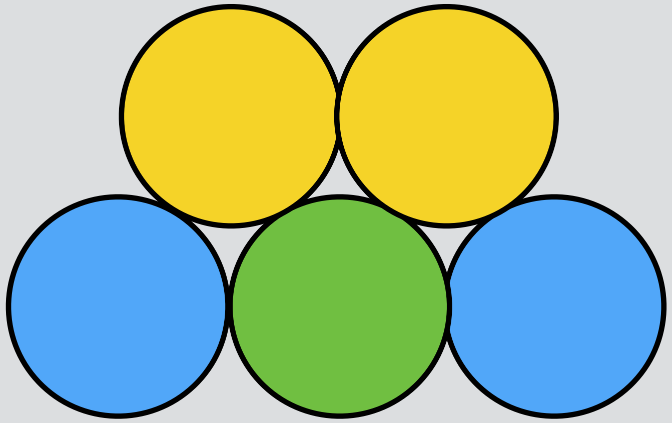


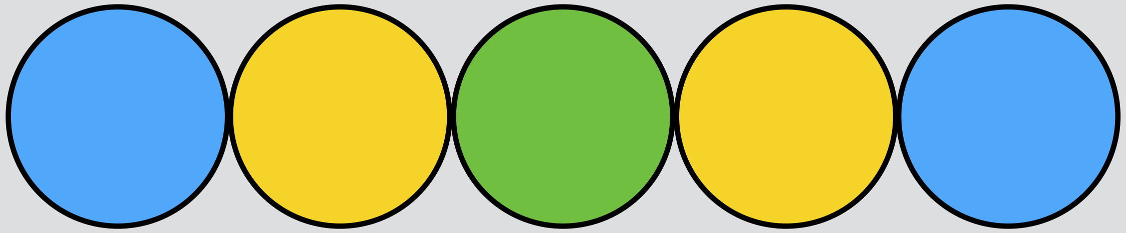


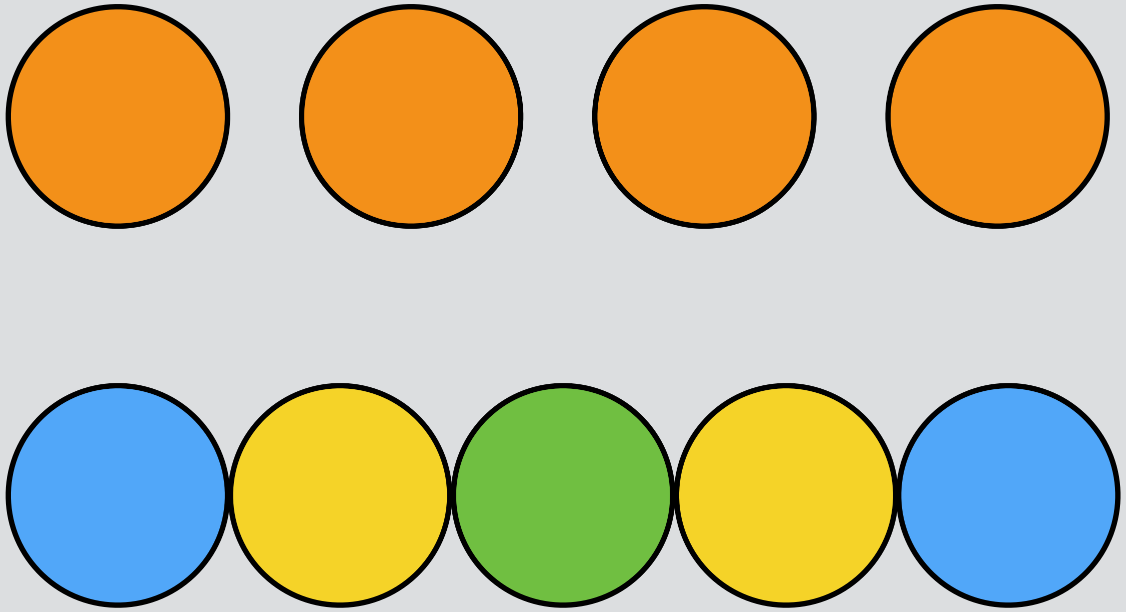


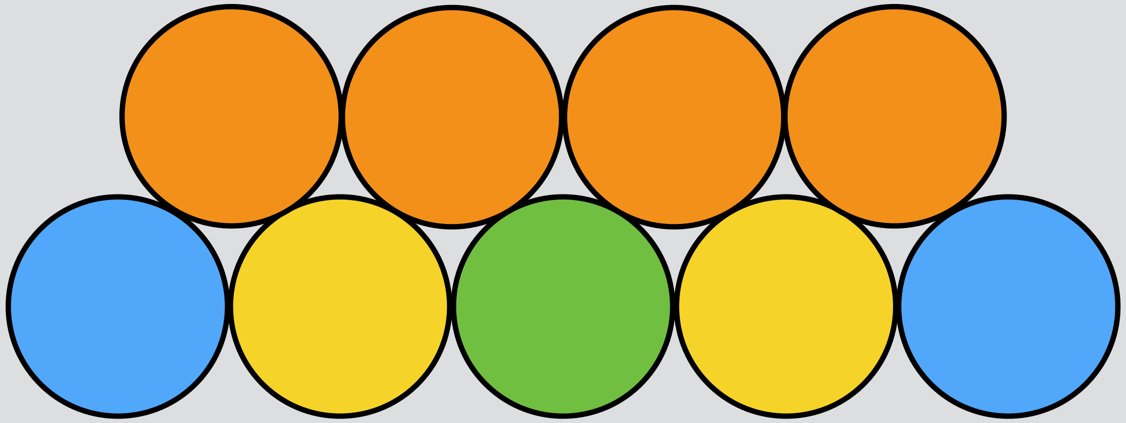


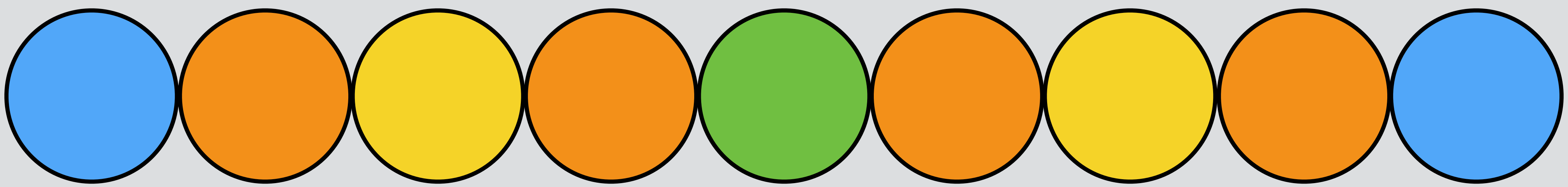


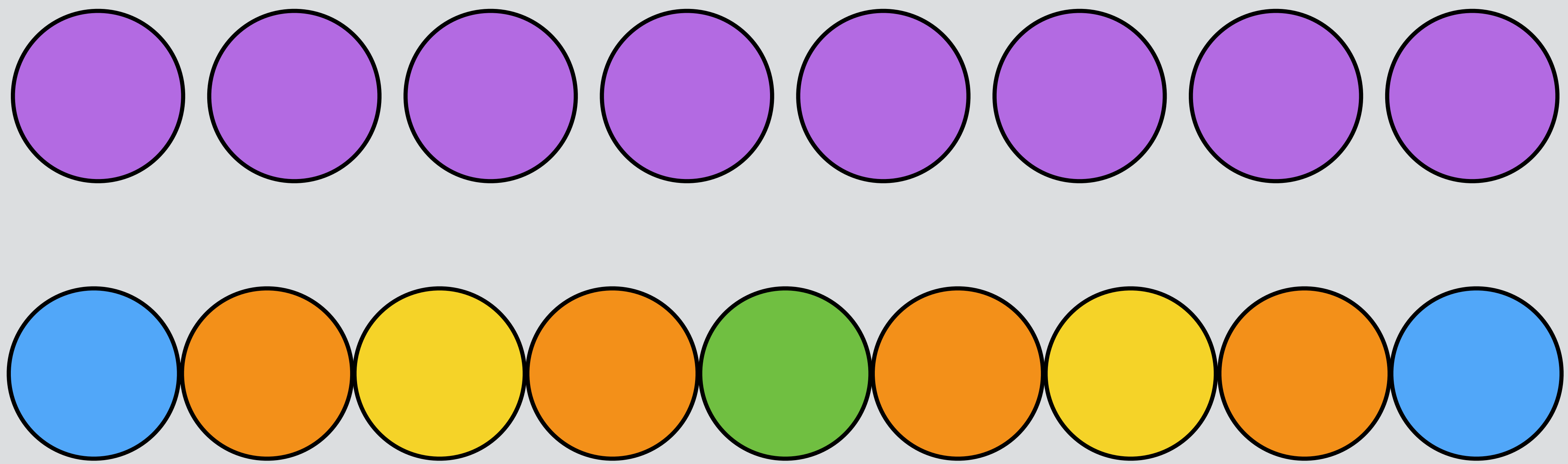


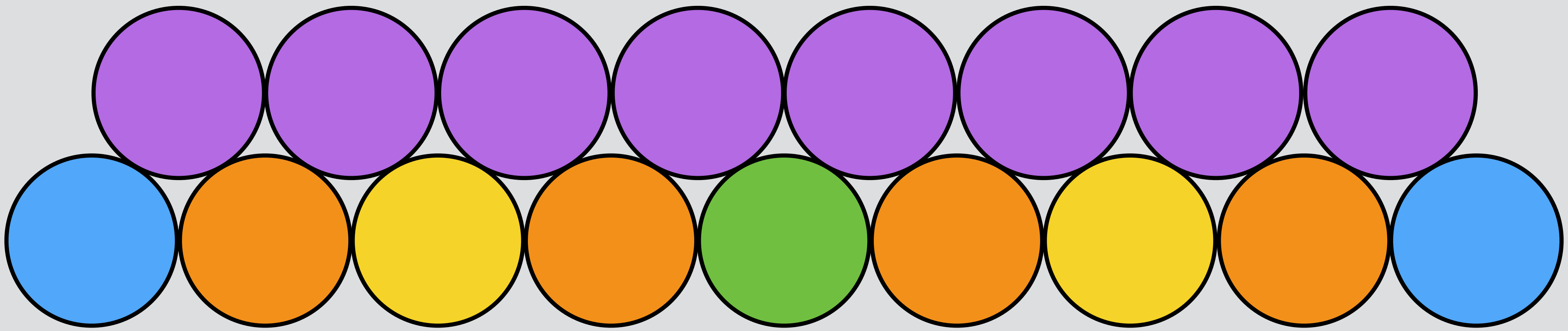


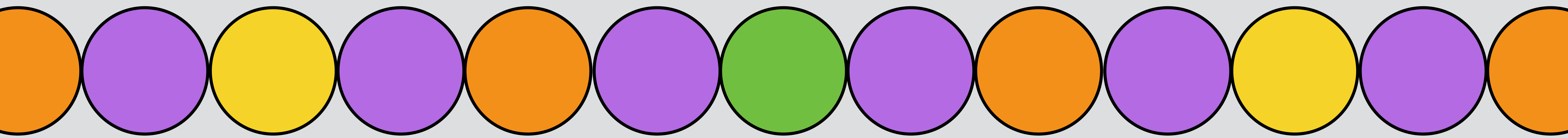




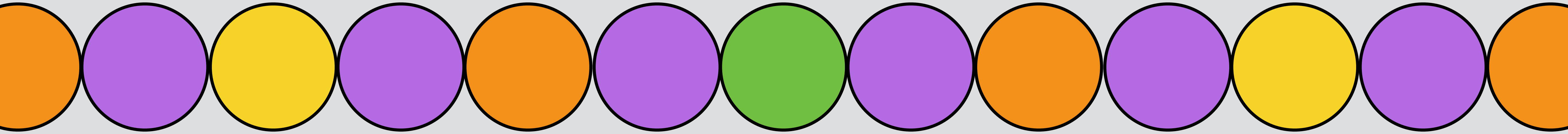






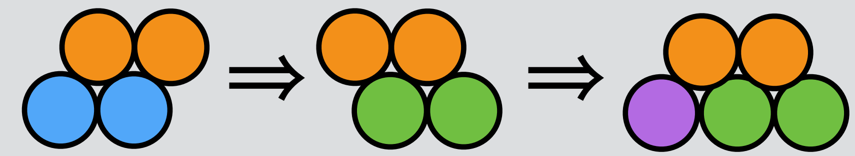


Exponential growth!

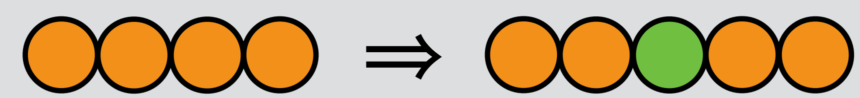


Active self-assembly models

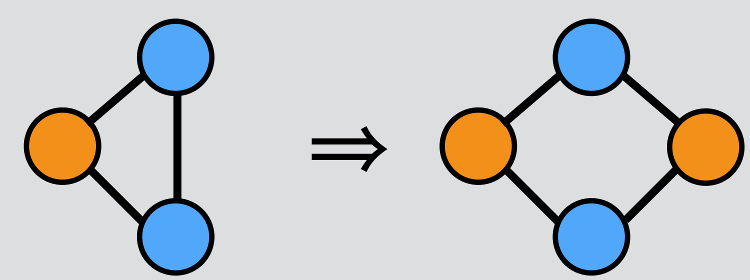
Nubots [Woods et al. ITCS 2012]:
2D, flexible and rigid bonds, stateful particles.



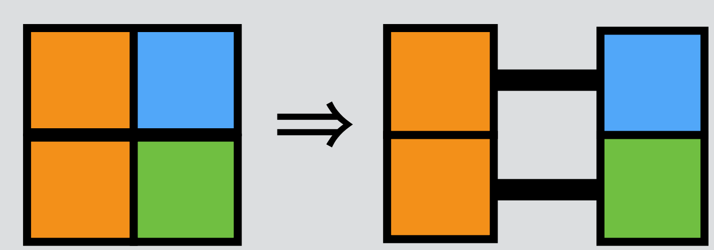
Insertion systems [Dabby, Chen SODA 2013]:
1D, fixed shape, stateless particles.



Graph grammars [Klavins et al. ICRA 2004]:
Geometry-less, stateless particles.



Crystalline robots [Rus, Vona ICRA 1999]:
3D, stateful particles, global communication.



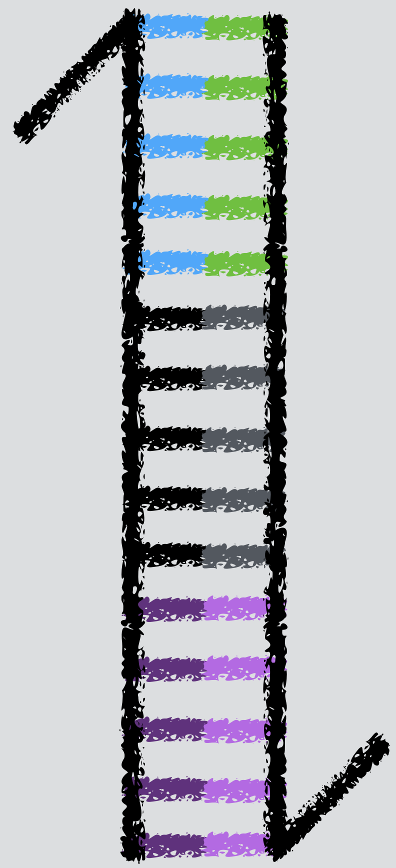
Insertion systems

- Introduced by [Dabby, Chen 2013].
- A model of active self-assembly.
    - Implementable in DNA.
    - Capable of exponential growth.
    - Grows a linear structure by insertion of particles.

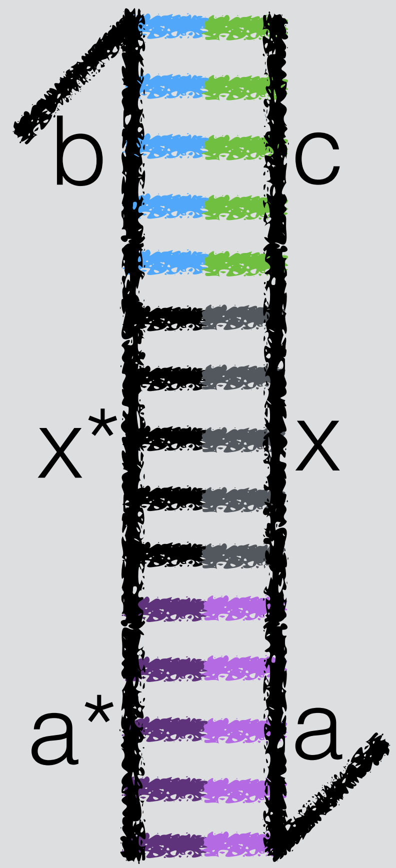
Insertion systems

- Introduced by [Dabby, Chen 2013].
- A model of active self-assembly.
    - Implementable in DNA.
    - Capable of exponential growth.
    - Grows a linear structure by insertion of particles.
- Our work: bound capabilities of insertion systems.

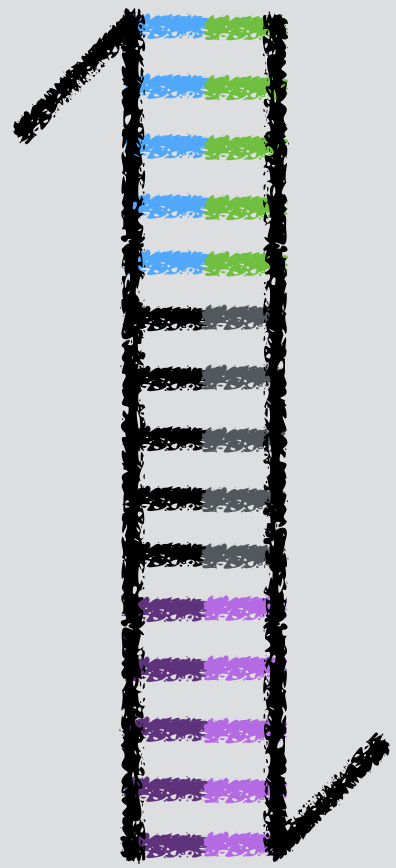
DNA insertions



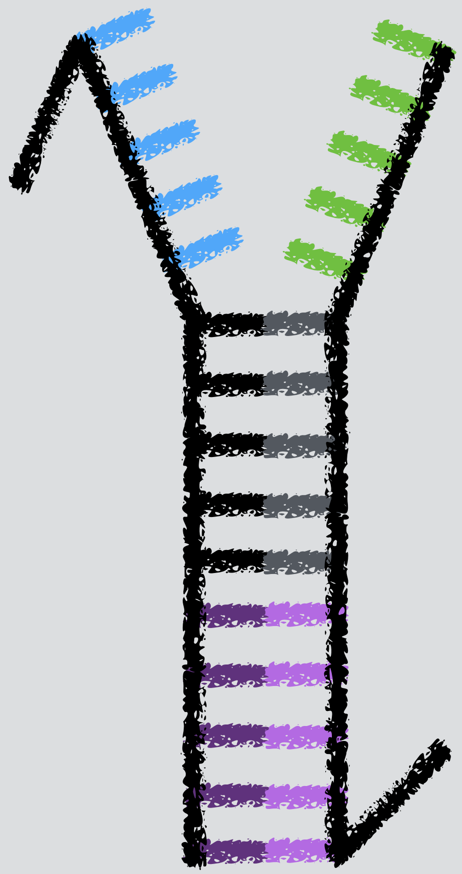
DNA insertions



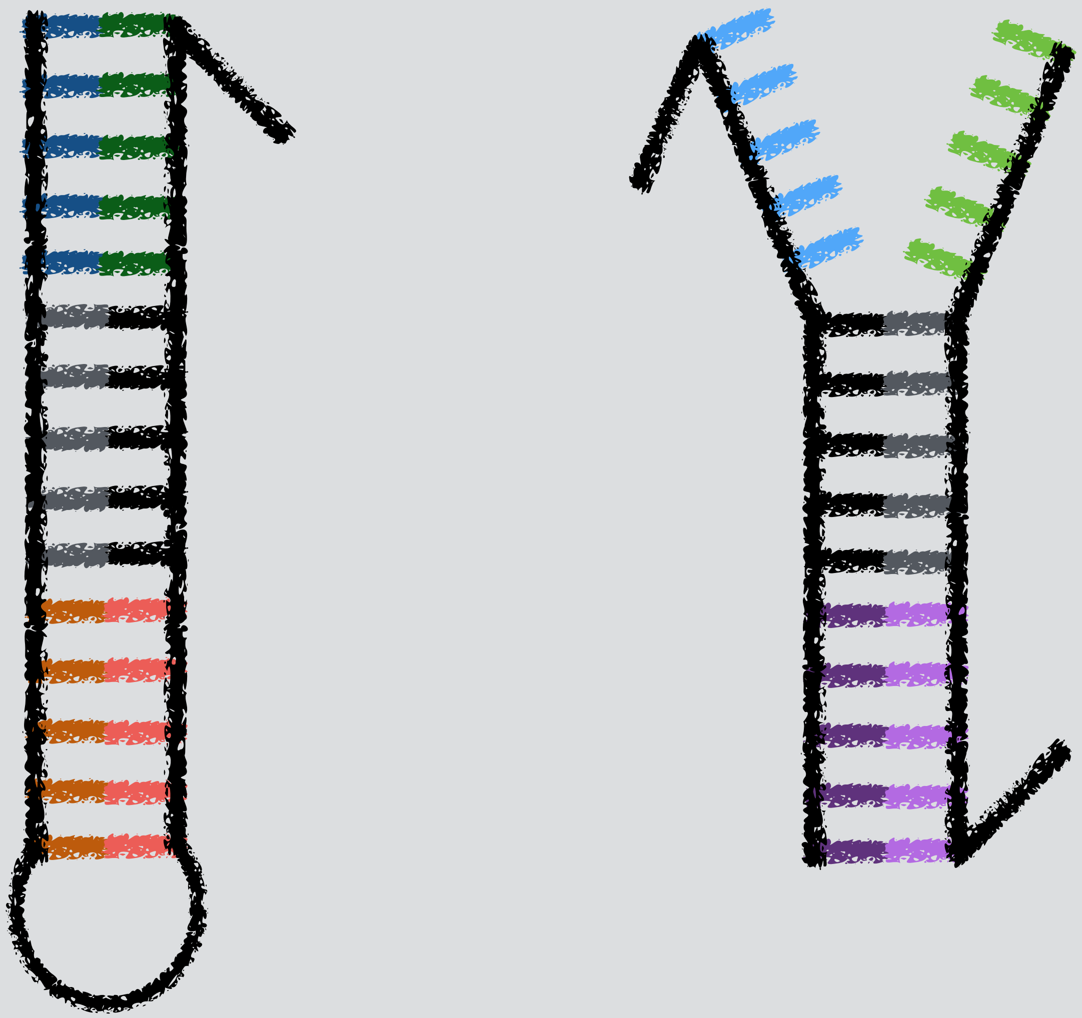
DNA insertions



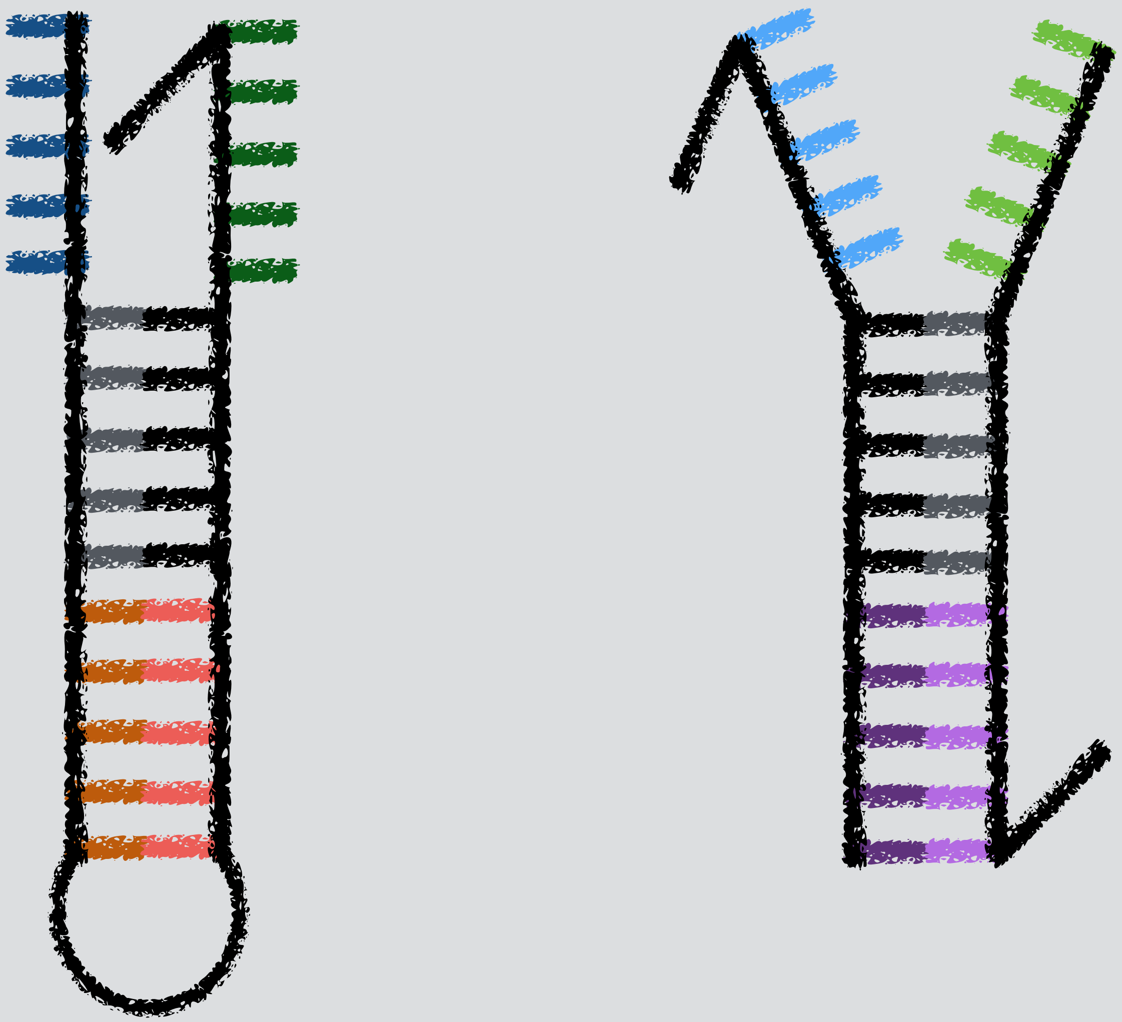
DNA insertions



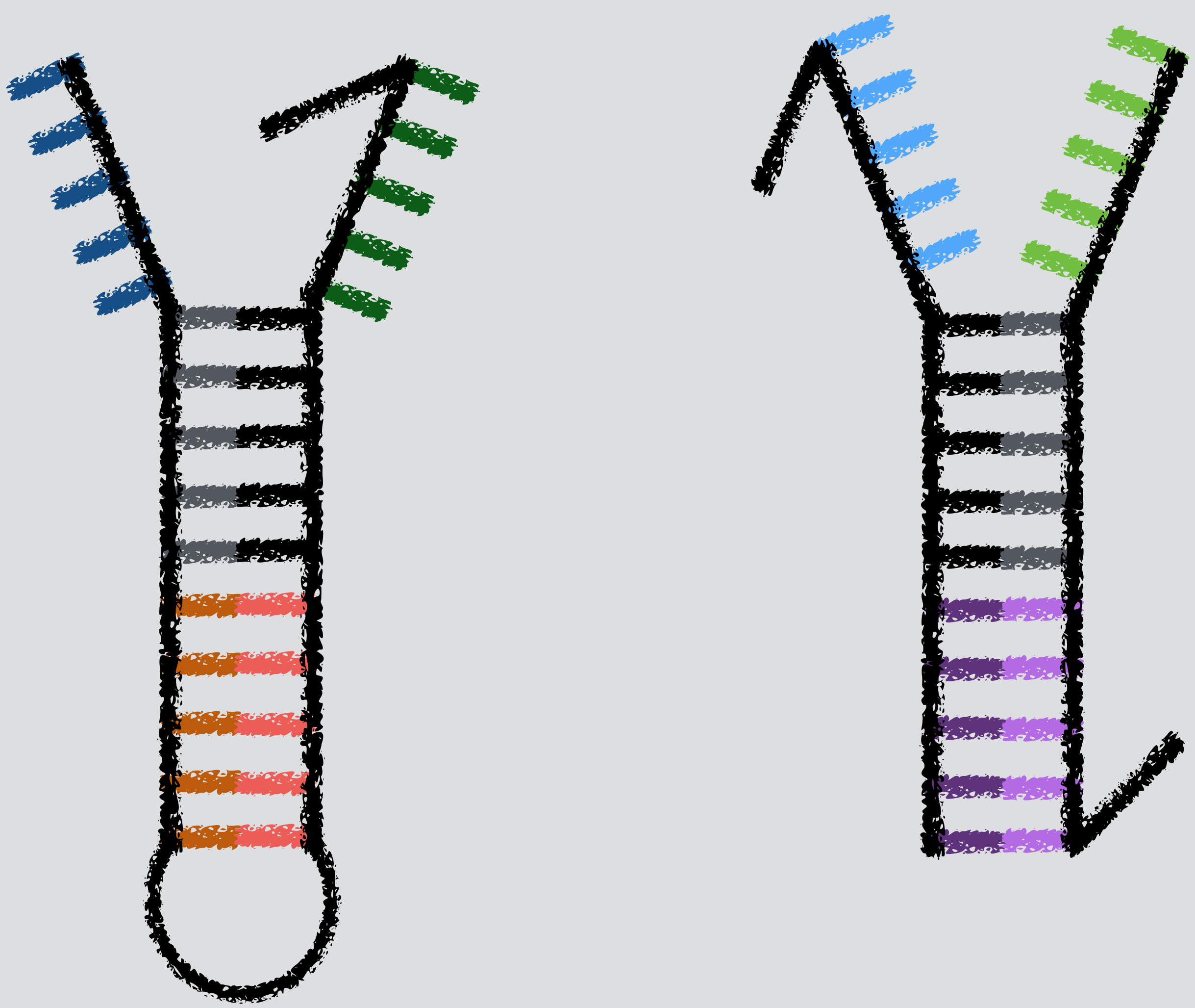
DNA insertions



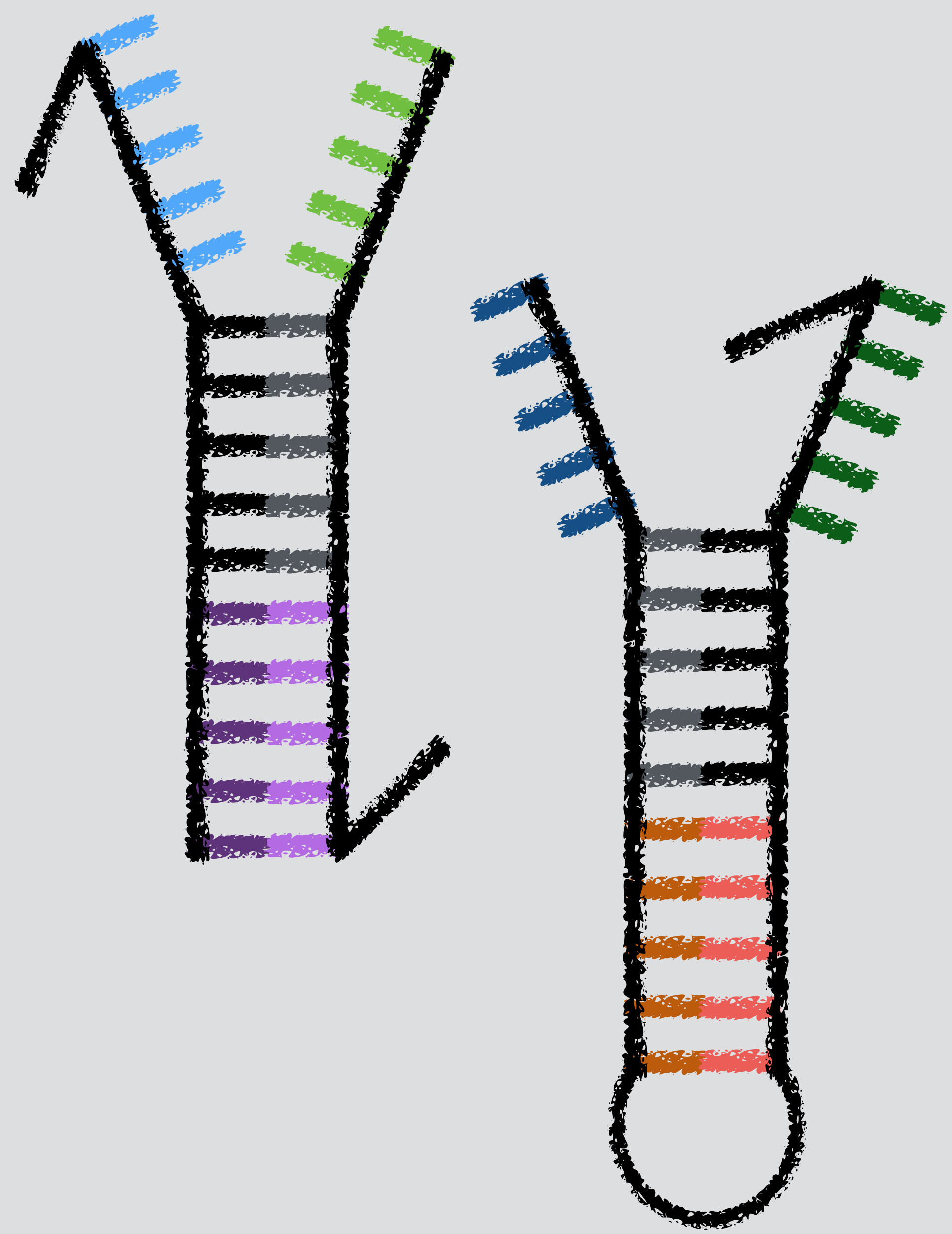
DNA insertions



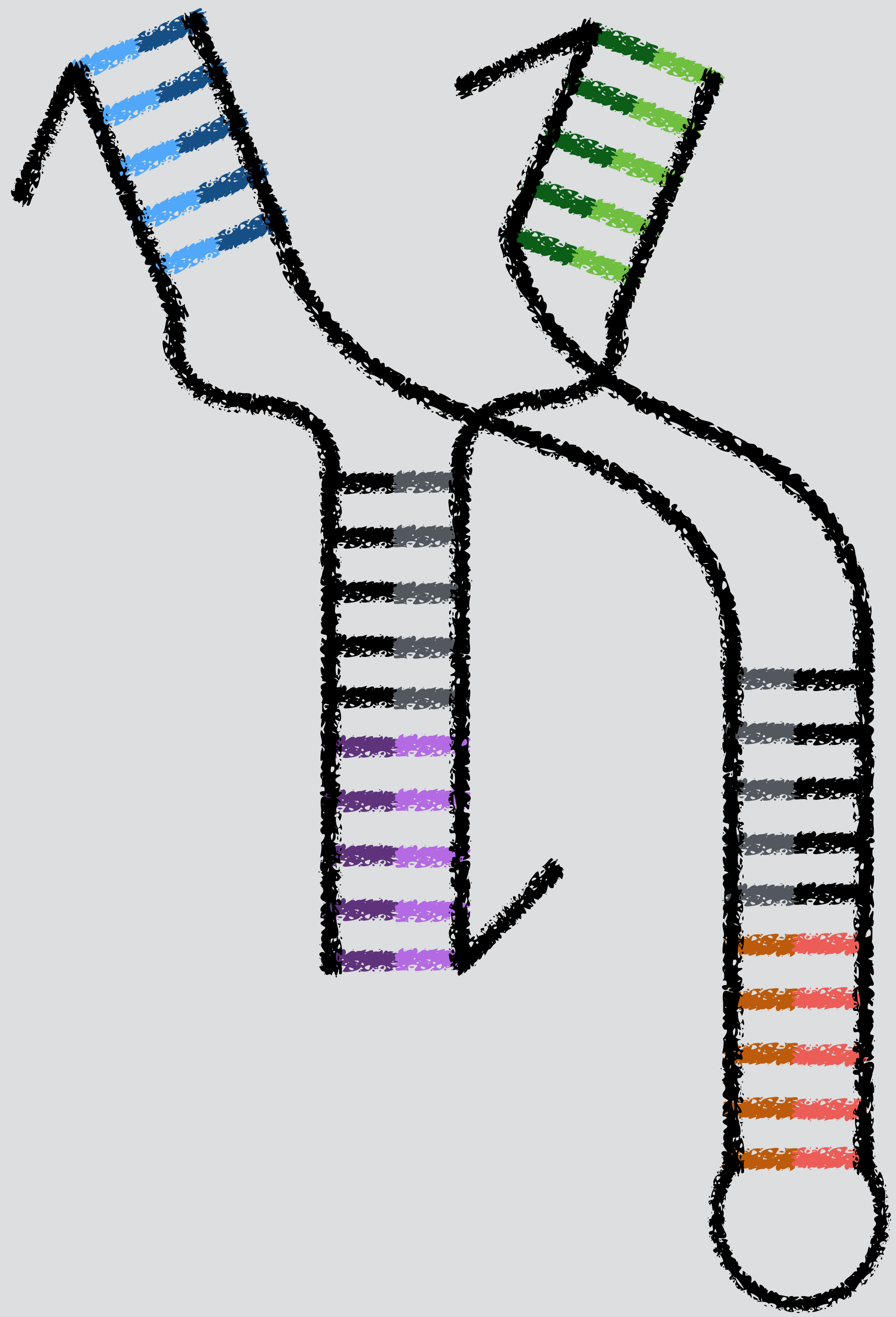
DNA insertions



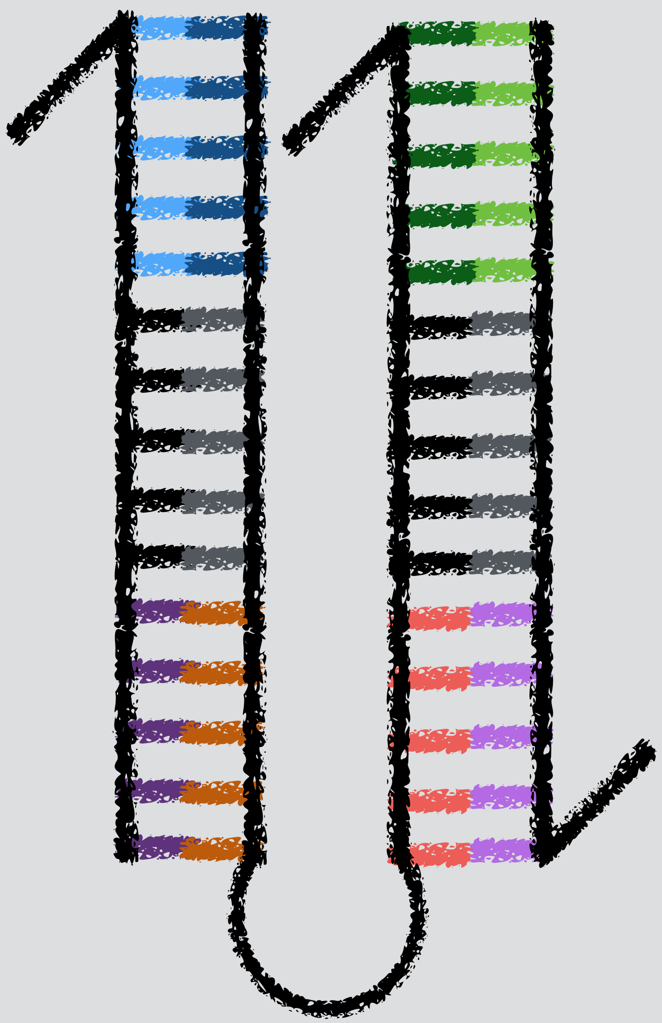
DNA insertions



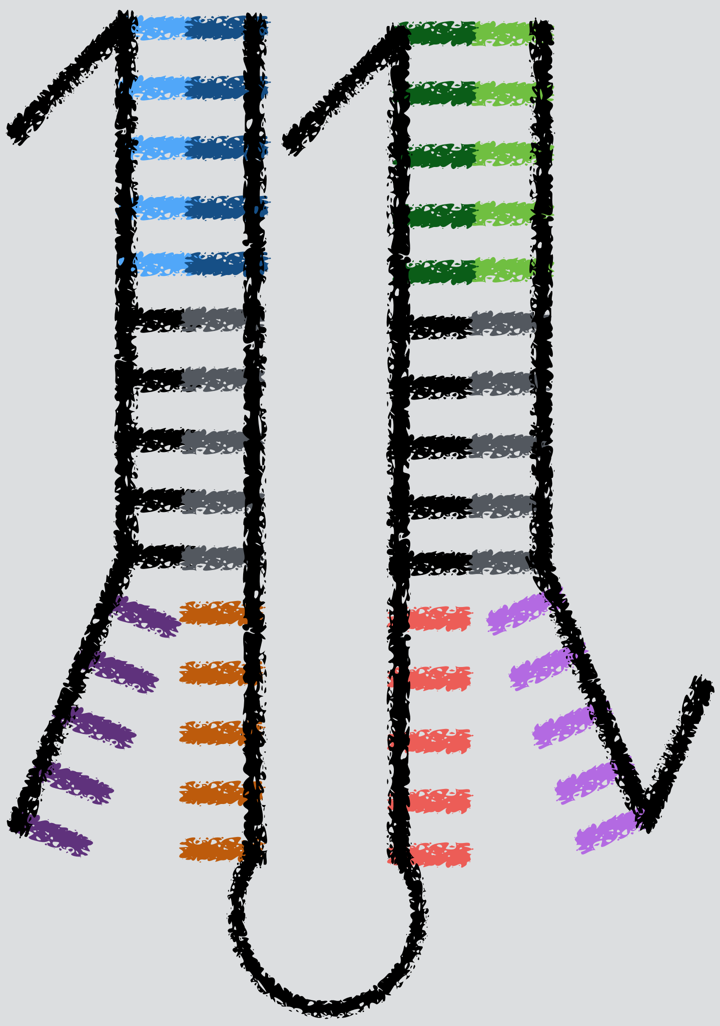
DNA insertions



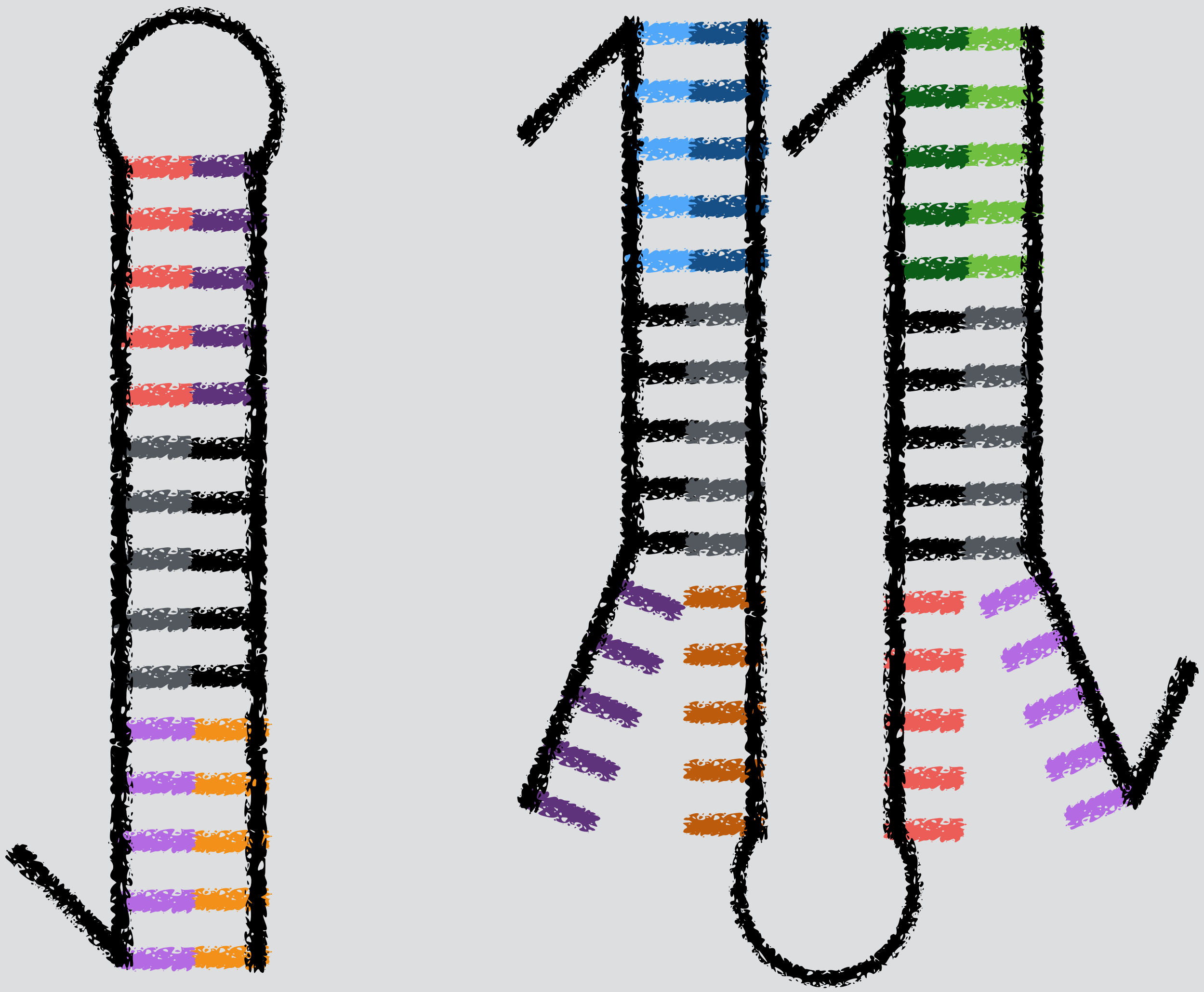
DNA insertions



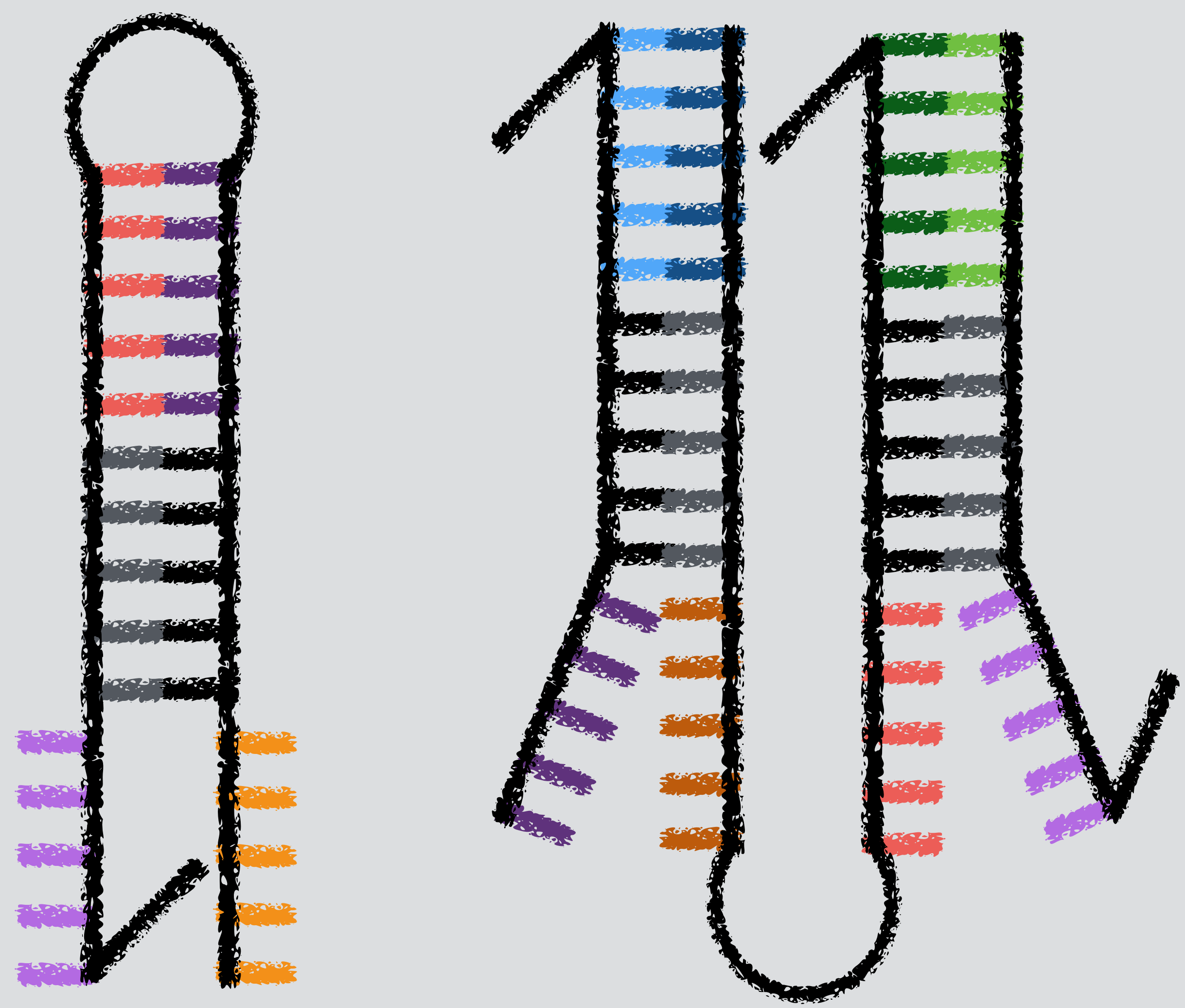
DNA insertions



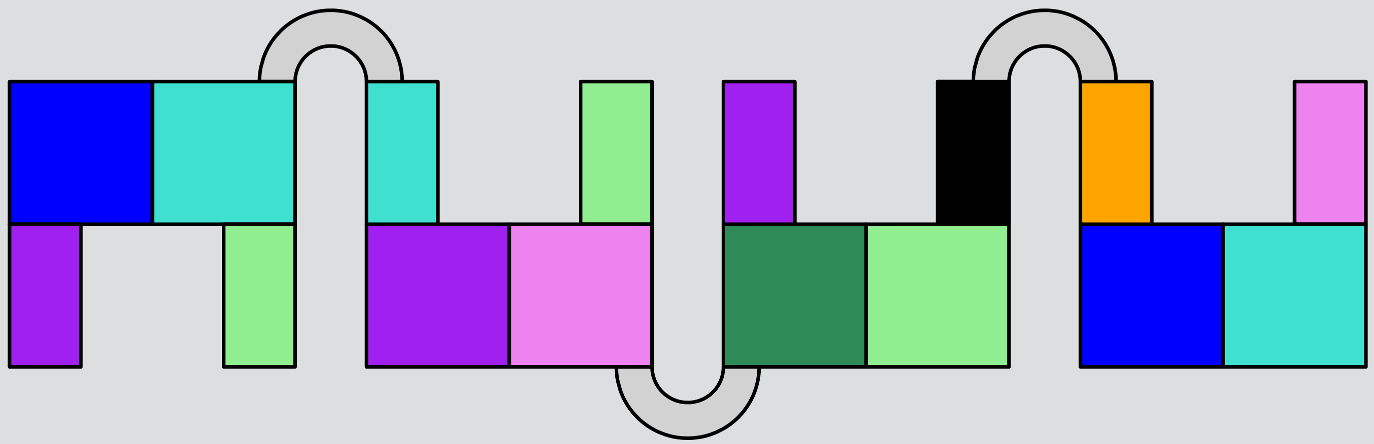
DNA insertions



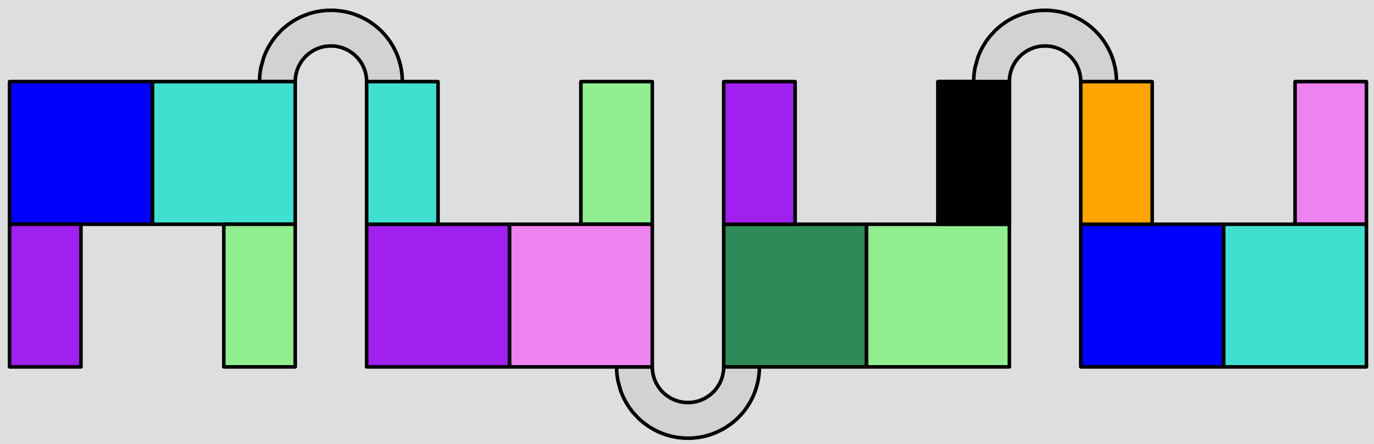
DNA insertions



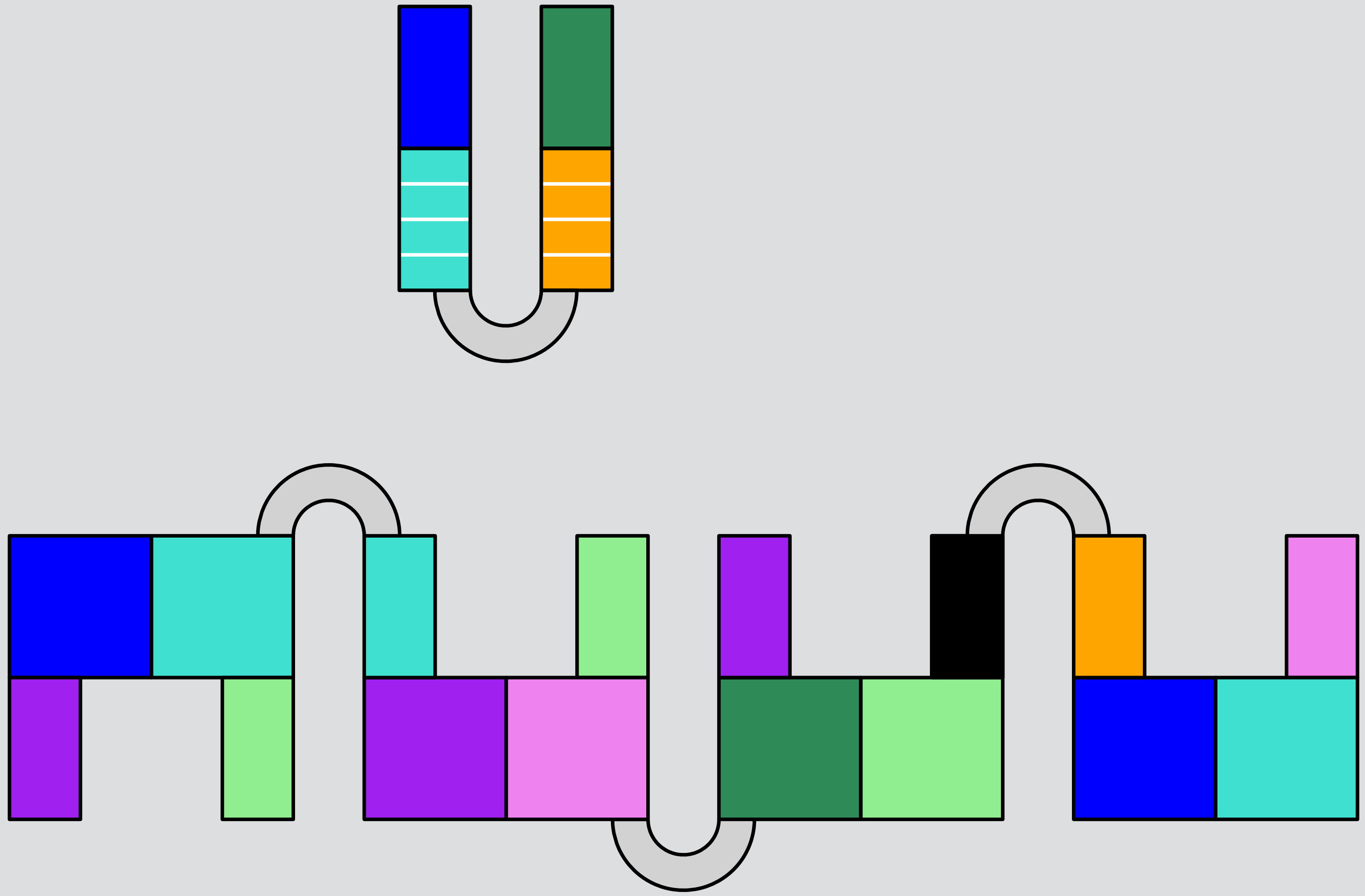
Insertion systems



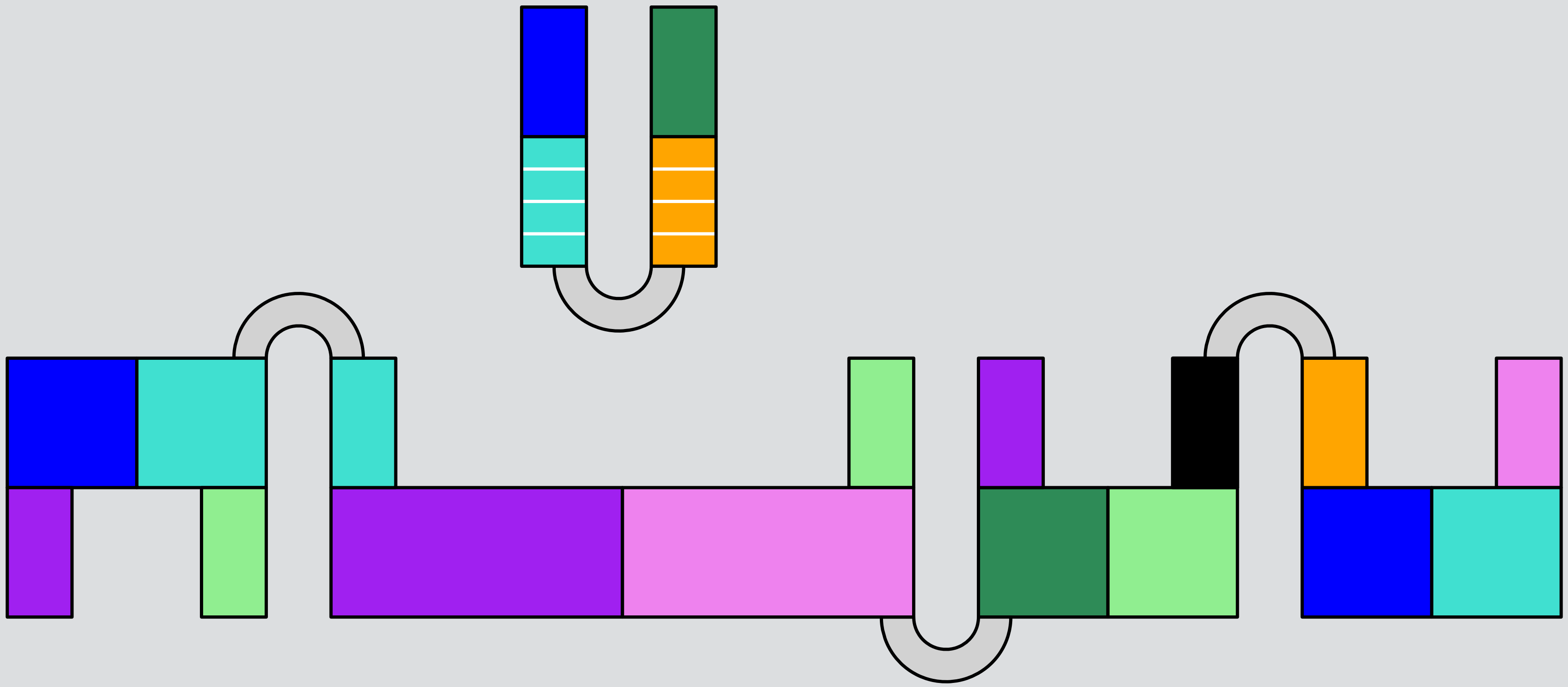
Insertion systems



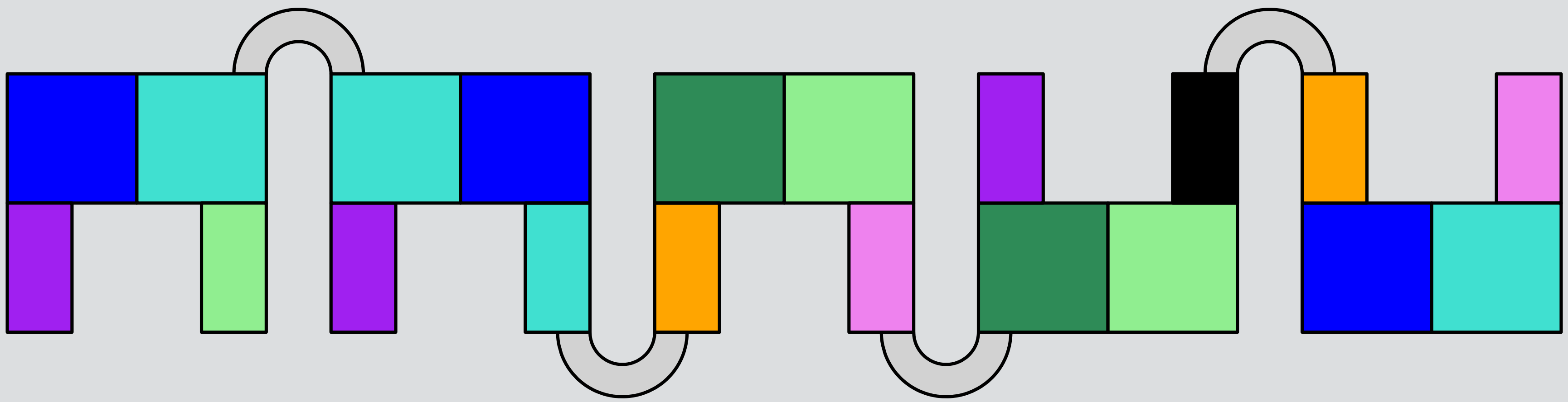
Insertion systems



Insertion systems

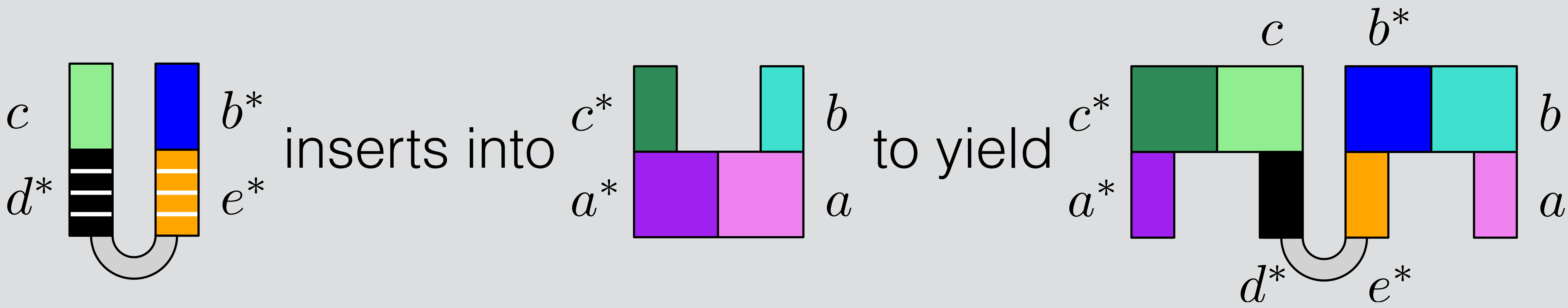


Insertion systems

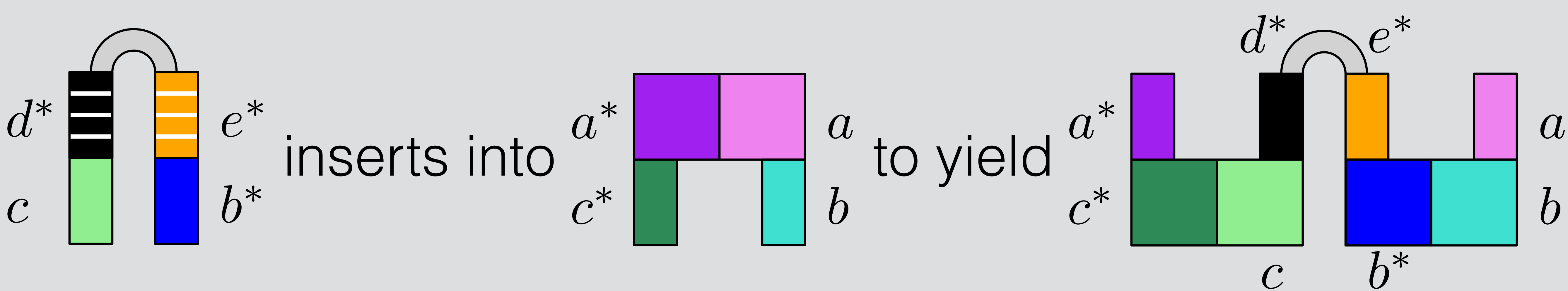


Definitions and examples

Insertions



$(c, d^*, e^*, b^*)^+$ inserts into $(a^*, c^*)(b, a)$ to yield $(a^*, c^*)(c, d^*, e^*, b^*)(b, a)$



$(d^*, c, b^*, e^*)^-$ inserts into $(c^*, a^*)(a, b)$ to yield $(c^*, a^*)(d^*, c, b^*, e^*)(a, b)$

Insertion system:

Monomer types: $(1,x,2,b)^{+}$ $(x,2^*,a,3^*)^{-}$ $(3,x,4,b)^{+}$

Initiator: $(a,1^*)(b^*,a^*)$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3^*)^-$ $(3,x,4,b)^+$

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3^*)^-$ $(3,x,4,b)^+$

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3^*)^-$ $(3,x,4,b)^+$

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3^*)^-$ $(3,x,4,b)^+$

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(3,x,4,b)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3^*)^-$ $(3,x,4,b)^+$

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(b^*,a^*)$$
$$\Downarrow$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(3,x,4,b)(b^*,a^*)$$

Terminal polymer of length 5

Insertion time

- Each monomer type has a concentration in [0,1].
- Concentrations of all types in a system must sum to ≤ 1.
- An insertion occurs after time t with:
  - t an exponential random variable with rate c.
  - c is the total concentration of insertable monomers.

Insertion system:

Monomer types: $(1,x,2,b)^{+}$ $(x,2^{*},a,3^{*})^{-}$ $(3,x,4,b)^{+}$

Initiator: $(a,1^{*})(b^{*},a^{*})$

Insertion system:

Monomer types:	$(1,x,2,b)^{+}$	$(x,2^*,a,3^*)^{-}$	$(3,x,4,b)^{+}$
Concentrations:	0.25	0.25	0.5

Initiator: $(a,1^*)(b^*,a^*)$

Insertion system:

Monomer types:	$(1,x,2,b)^+$	$(x,2^*,a,3^*)^-$	$(3,x,4,b)^+$
Concentrations:	0.25	0.25	0.5

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^{+}$ $(x,2^{*},a,3^{*})^{-}$ $(3,x,4,b)^{+}$

Concentrations: 0.25 0.25 0.5

Initiator: $(a,1^{*})(b^{*},a^{*})$

$$(a,1^{*})(b^{*},a^{*})$$
$$\Downarrow t_1$$
$$(a,1^{*})(1,x,2,b)(b^{*},a^{*})$$
$$\Downarrow t_2$$
$$(a,1^{*})(1,x,2,b)(x,2^{*},a,3^{*})(b^{*},a^{*})$$
$$\Downarrow t_3$$
$$(a,1^{*})(1,x,2,b)(x,2^{*},a,3^{*})(3,x,4,b)(b^{*},a^{*})$$

Terminal polymer of length 5

Insertion system:

Monomer types:	$(1,x,2,b)^+$	$(x,2^*,a,3^*)^-$	$(3,x,4,b)^+$
Concentrations:	0.25	0.25	0.5

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(b^*,a^*)$$
$$\Downarrow t_3$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,3^*)(3,x,4,b)(b^*,a^*)$$

Terminal polymer of length 5

Expected time: $t_1 + t_2 + t_3$, with

$E[t_1] = E[t_2] = 4$, $E[t_3] = 2$.

$4 + 4 + 2 = 12$

Insertion system:

Monomer types:	$(1^*,2,2,1^*)^+$	$(x,0^*,2^*,x)^-$	$(x,2^*,0,x)^-$
Concentrations:	0.5	0.1	0.4

Initiator: $(0,1)(1,0^*)$

$(0,1)(1,0^*)$

Insertion system:

Monomer types:	$(1^*,2,2,1^*)^+$	$(x,0^*,2^*,x)^-$	$(x,2^*,0,x)^-$
Concentrations:	0.5	0.1	0.4

Initiator: $(0,1)(1,0^*)$

$$(0,1)(1,0^*)$$
$$\Downarrow t_1$$
$$(0,1)(1^*,2,2,1^*)(1,0^*)$$

Insertion system:

Monomer types: $(1^*,2,2,1^*)^+$ $(x,0^*,2^*,x)^-$ $(x,2^*,0,x)^-$

Concentrations: 0.5 0.1 0.4

Initiator: $(0,1)(1,0^*)$

$$(0,1)(1,0^*)$$
$$\Downarrow t_1$$
$$(0,1)(1^*,2,2,1^*)(1,0^*)$$
$$\Downarrow t_2$$
$$(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(1,0^*)$$

Insertion system:

Monomer types: $(1^*,2,2,1^*)^+$ $(x,0^*,2^*,x)^-$ $(x,2^*,0,x)^-$

Concentrations: 0.5 0.1 0.4

Initiator: $(0,1)(1,0^*)$

$$
\begin{array}{c}
(0,1)(1,0^*) \\
\Downarrow t_1 \\
(0,1)(1^*,2,2,1^*)(1,0^*) \\
\Downarrow t_2 \\
(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(1,0^*) \\
\Downarrow t_3 \\
(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(x,2^*,0,x)(1,0^*)
\end{array}
$$

Insertion system:

Monomer types: $(1^*,2,2,1^*)^+$ $(x,0^*,2^*,x)^-$ $(x,2^*,0,x)^-$

Concentrations: 0.5 0.1 0.4

Initiator: $(0,1)(1,0^*)$

$$(0,1)(1,0^*)$$
$$\Downarrow t_1$$
$$(0,1)(1^*,2,2,1^*)(1,0^*)$$
$$\Downarrow t_2$$
$$(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(1,0^*)$$
$$\Downarrow t_3$$
$$(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(x,2^*,0,x)(1,0^*)$$

Terminal polymer of length 5

Insertion system:

Monomer types: $(1^*,2,2,1^*)^+$ $(x,0^*,2^*,x)^-$ $(x,2^*,0,x)^-$

Concentrations: 0.5 0.1 0.4

Initiator: $(0,1)(1,0^*)$

$$
\begin{aligned}
&(0,1)(1,0^*)\\
&\Downarrow t_1\\
&(0,1)(1^*,2,2,1^*)(1,0^*)\\
&\Downarrow t_2\\
&(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(1,0^*)\\
&\Downarrow t_3\\
&(0,1)(x,0^*,2^*,x)(1^*,2,2,1^*)(x,2^*,0,x)(1,0^*)
\end{aligned}
$$

Terminal polymer of length 5

Expected time: $t_1 + \max(t_2, t_3)$, with

$E[t_1] = 2$, $E[t_2] = 10$, $E[t_3] = 2.5.$

$2 + 10.5 = 12.5$

Insertion system:

Monomer types:	$(1,x,2,b)^{+}$	$(x,2^*,a,1^*)^{-}$	$(x,2^*,a,x)^{-}$
Concentrations:	0.5	0.4	0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$

Insertion system:

Monomer types:	$(1,x,2,b)^+$	$(x,2^*,a,1^*)^-$	$(x,2^*,a,x)^-$
Concentrations:	0.5	0.4	0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$$

Expected time: $t_1 + t_2$*, with*

$E[t_1] = E[t_2] = 2.$

$2 + 2 = 4$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$
\begin{gathered}
(a,1^*)(b^*,a^*) \\
\Downarrow t_1 \\
(a,1^*)(1,x,2,b)(b^*,a^*) \\
\Downarrow t_2 \\
(a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)
\end{gathered}
$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(b^*,a^*)$$
$$\Downarrow t_3$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(b^*,a^*)$$

Insertion system:

Monomer types: $(1,x,2,b)^{+}$ $(x,2^*,a,1^*)^{-}$ $(x,2^*,a,x)^{-}$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$
\begin{gathered}
(a,1^*)(b^*,a^*)\\
\Downarrow t_1\\
(a,1^*)(1,x,2,b)(b^*,a^*)\\
\Downarrow t_2\\
(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(b^*,a^*)\\
\Downarrow t_3\\
(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(b^*,a^*)\\
\Downarrow t_4\\
(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)
\end{gathered}
$$

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$$(a,1^*)(b^*,a^*)$$
$$\Downarrow t_1$$
$$(a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_2$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(b^*,a^*)$$
$$\Downarrow t_3$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(b^*,a^*)$$
$$\Downarrow t_4$$
$$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$$

Expected time: $t_1 + t_2 + t_3 + t_4$*, with*

$E[t_1] = E[t_2] = E[t_3] = E[t_4] = 2.$

$2 + 2 + 2 + 2 = 8$

Expressive power

Expressive Power

- Insertion systems: initiator + set of monomers = set of polymers, with terminal polymer subset.
- Context-free grammars: start symbol + set of rules = set of partial derivations, with string subset.
- Do insertion systems and context-free grammars have equal “expressive-ness”?

Expressive Power

- Insertion systems: initiator + set of monomers = set of polymers, with terminal polymer subset.

- Context-free grammars: start symbol + set of rules = set of partial derivations, with string subset.

- Do insertion systems and context-free grammars have equal “expressive-ness”? *Yes.*

Expressive Power

Theorem: every insertion system can be expressed as a context-free grammar. [Dabby, Chen 2013]

Theorem: every context-free grammar can be expressed as an insertion system. [This work]

Expressive Power

Theorem: every insertion system can be expressed as a context-free grammar. [Dabby, Chen 2013]

Theorem: every context-free grammar can be expressed as an insertion system. [This work]

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,3)^-$ $(3,x,4,b)^+$

Concentrations: 0.25 0.25 0.25

Initiator: $(a,1)(b^*,a^*)$

Context-free grammar:

Production rules:

$S_{(a,1)(b^*,a^*)} \to S_{(a,1)(1,x)} S_{(2,b)(b^*,a^*)}$

$S_{(2,b)(b^*,a^*)} \to S_{(2,b)(x,2^*)} S_{(a,3)(b^*,a^*)}$

$S_{(a,3)(b^*,a^*)} \to S_{(a,3)(3,x)} S_{(4,b)(b^*,a^*)}$

$S_{(a,1)(1,x)} \to (a, 1)(1, x)$

$S_{(2,b)(x,2^*)} \to (2,b)(x,2^*)$

$S_{(a,3)(3,x)} \to (a,3)(3,x)$

$S_{(4,b)(x,4^*)} \to (4,b)(x,4^*)$

Start symbol: $S_{(a,1)(b^*,a^*)}$

Expressive Power

<u>Theorem:</u> every insertion system can be expressed as a context-free grammar. [Dabby, Chen 2013]

<u>Theorem:</u> every context-free grammar can be expressed as an insertion system. [This work]

Expressive Power

Theorem: every insertion system can be expressed as a context-free grammar. [Dabby, Chen 2013]

Theorem: every context-free grammar can be expressed as an insertion system. [This work]

Turning rules into monomers

Rule: $A \rightarrow B\ C$

Derivation step: $cdADee \Rightarrow cdBCDee$

Monomer type: $(1,x,2,b)^{+}$

Insertion: $(a,1)(b^{*},a^{*}) \Rightarrow (a,1)(1,x,2,b)(b^{*},a^{*})$

Rules completely replace non-terminals.
Insertions do not completely replace insertion sites.

Turning rules into monomers

Rule: $A_2 \rightarrow A_4 A_5$

Derivation step: $A_1 A_2 A_3 \Rightarrow A_1 A_4 A_5 A_3$

(n non-terminals total)

Insertion site: $u^*, s_a^*)(s_d, u$ with $a+d \bmod n = 2$

Site modification: $u^*, s_a^*)(s_d, u$

$$\Downarrow$$

$$\vdots$$

$$\Downarrow$$

$u^*, s_a^*)(s_b, u, \ldots, u^*, s_c^*)(s_d, u$

with $a+b \bmod n = 4$, $c+d \bmod n = 5$

Monomer types: $\Theta(n)$ per rule.

Polymer lengths

Insertion system:

Monomer types: $(1,x,2,b)^{+}$ $(x,2^{*},a,1^{*})^{-}$ $(x,2^{*},a,x)^{-}$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^{*})(b^{*},a^{*})$

$(a,1^{*})(1,x,2,b)(x,2^{*},a,x)(b^{*},a^{*})$

$(a,1^{*})(1,x,2,b)(x,2^{*},a,1^{*})(1,x,2,b)(x,2^{*},a,x)(b^{*},a^{*})$

$(a,1^{*})(1,x,2,b)(x,2^{*},a,1^{*})(1,x,2,b)(x,2^{*},a,1^{*})(1,x,2,b)(x,2^{*},a,x)(b^{*},a^{*})$

•••

Context-free grammar:

Production rules: $S \rightarrow A$ $A \rightarrow aaA$ $A \rightarrow aa$

Start symbol: S

aa

aaaa

aaaaaa

•••

Insertion system:

Monomer types: $(1,x,2,b)^+$ $(x,2^*,a,1^*)^-$ $(x,2^*,a,x)^-$

Concentrations: 0.5 0.4 0.1

Initiator: $(a,1^*)(b^*,a^*)$

$(a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$

$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$

$(a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(x,2^*,a,1^*)(1,x,2,b)(x,2^*,a,x)(b^*,a^*)$

•••

Constructing arbitrarily long polymers is easy if infinite polymers allowed.

Context-free grammar:

Production rules: $S \rightarrow A$ $A \rightarrow aaA$ $A \rightarrow aa$

Start symbol: S

aa

aaaa

aaaaaa

•••

Constructing long polymers

Theorem: a system with k monomer types constructing a finite number of polymers can construct:

- polymers of length $2^{\Theta(k^{1/2})}$ [Dabby, Chen 2013]
- polymers of length $2^{\Theta(k^{3/2})}$ [This work]
- only polymers of length $2^{O(k^{3/2})}$ [This work]

Constructing long polymers

Theorem: a system with k monomer types constructing a finite number of polymers can construct:

- polymers of length $2^{\Theta(k^{1/2})}$ [Dabby, Chen 2013]
- polymers of length $2^{\Theta(k^{3/2})}$ [This work]
- only polymers of length $2^{O(k^{3/2})}$ [This work]

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.

$$\ldots u,0)(0,u^*\ldots$$

$$\Downarrow$$

$$\ldots u,0)(0^*,1,x,0^*)(0,u^*\ldots$$

$$\Downarrow$$

$$\ldots u,0)(x,u^*,1^*,1^*)(0^*,1,x,0^*)(0,u^*\ldots$$

$$\Downarrow$$

$$\ldots u,0)(x,u^*,1^*,1^*)(1,\ x,\ 2,\ 0)(0^*,1,x,0^*)(0,u^*\ldots$$

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.

…u,0)(0,u*…

⇓

…u,0)(0*,1…

⇓

…1*,1*)(0*,1…

⇓

…2, 0)(0*,1…

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.

$$
\begin{gathered}
\ldots u,0)(0,u^{*}\ldots \\
\Downarrow \\
\ldots u,0)(0^{*},1\ldots \\
\Downarrow \\
\ldots 1^{*},1^{*})(0^{*},1\ldots \\
\Downarrow \\
\ldots 2,\ 0)(0^{*},1\ldots
\end{gathered}
$$

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.
- Ingredient 1: long insertion sequence with no repeated insertion sites.

…0,0)(0,0*…

$\Downarrow$ O(1)

…0,0)(1,0*…

$\Downarrow$ O(1)

⋮

$\Downarrow$ O(1)

…r, r)(r,r*…

$\Theta(k^{3/2})$

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.

- Ingredient 1: long insertion sequence with no repeated insertion sites.

- Ingredient 2: duplication of each site in sequence.

$$\ldots 0,0)(0,0^*\ldots$$
$$\Downarrow O(1)$$
$$\ldots 0,0)(1,0^*\ldots 0,0)(1,0^*\ldots$$
$$\Downarrow O(1)$$
$$\ldots 0,\ 0)(2,0^*\ldots 0,0)(2,0^*\ldots 0,0)(2,0^*\ldots 0,0)(2,0^*\ldots$$

Constructing long polymers

- Consider *insertion sequences*: repeated insertions into the site resulting from previous insertion.
- Ingredient 1: long insertion sequence
  with no repeated insertion sites.
- Ingredient 2: duplication of each site in sequence.
- Combine these for long polymers.

Constructing long polymers

…0,0)(0,0*…

$\Downarrow$ O(1)

…0,0)(1,0*…0,0)(1,0*…

$\Downarrow$ O(1)

…0, 0)(2,0*…0,0)(2,0*…0,0)(2,0*…0,0)(2,0*…

$\Downarrow$ O(1)

⋮

$\Downarrow$ O(1)

…r, r)(r,r*…r, r)(r,r*…r, r)(r,r*…r,r)(r,r*…r,r)(r,r*…r,r)(r,r*…

$\Theta(k^{3/2})$

$2^{\Theta(k^{3/2})}$

Constructing long polymers

- Ingredient 1: long insertion sequence with no repeated insertion sites.

$$s_a, s_b)(s_c, s_a{}^*$$

variables

Use $r+1 = \Theta(k^{1/2})$ values of a, b, c $\Rightarrow \Theta(r^3) = \Theta(k^{3/2})$ insertion sites.

Case:	$b < r$	$b = r,\ c < r$	$b = c = r,\ a < r$
Insertion sequence:	$s_a, s_b)(s_c, s_a{}^* \overset{O(1)}{\Downarrow} s_a, s_{b+1})(s_c, s_a{}^*$	$s_a, s_r)(s_c, s_a{}^* \overset{O(1)}{\Downarrow} s_a, s_0)(s_{c+1}, s_a{}^*$	$s_a, s_r)(s_r, s_a{}^* \overset{O(1)}{\Downarrow} s_{a+1}, s_0)(s_0, s_{a+1}{}^*$
Result:	$++b$	$b = 0,\ ++c$	$b = c = 0,\ ++a$

Constructing long polymers

- Ingredient 1: long insertion sequence with no repeated insertion sites.

$$s_a, s_b)(s_c, s_a{}^*$$

variables

Use $r+1 = \Theta(k^{1/2})$ values of a, b, c $\Rightarrow \Theta(r^3) = \Theta(k^{3/2})$ insertion sites.

Case:	$b < r$	$b = r,\ c < r$	$b = c = r,\ a < r$
Insertion sequence:	$s_a, s_b)(s_c, s_a{}^* \Downarrow_{O(1)} s_a, s_{b+1})(s_c, s_a{}^*$	$s_a, s_r)(s_c, s_a{}^* \Downarrow_{O(1)} s_a, s_0)(s_{c+1}, s_a{}^*$	$s_a, s_r)(s_r, s_a{}^* \Downarrow_{O(1)} s_{a+1}, s_0)(s_0, s_{a+1}{}^*$
Result:	++b	b = 0, ++c	b = c = 0, ++a

Constructing long polymers

$\ldots s_a, s_r)(s_c, s_a{}^* \ldots$

Case:	$b = r,\ c < r$
Insertion sequence:	$s_a, s_r)(s_c, s_a{}^* \Downarrow_{O(1)} s_a, s_0)(s_{c+1}, s_a{}^*$
Result:	$b = 0,\ {+}{+}c$

$\ldots s_a, s_0{}^*)(s_{c+1}{}^*,\ s_a{}^* \ldots$

Constructing long polymers

Case:	$b = r,\ c < r$
Insertion sequence:	$s_a, s_r)(s_c, s_a^* \overset{O(1)}{\Downarrow} s_a, s_0)(s_{c+1}, s_a^*$
Result:	$b = 0,\ {+}{+}c$

$$
\begin{array}{c}
\ldots s_a, s_r)(s_c, s_a^* \ldots \\
\Downarrow \\
\ldots s_a, s_r)(s_r^*, s_{f(c)} \ldots \\
\Downarrow \\
\ldots s_a, s_r)(s_{f'(c)},\ s_a^* \ldots \\
\Downarrow \\
\ldots s_{f''(c)}, s_{f'(c)}^*)(s_{f'(c)},\ s_a^* \ldots \\
\Downarrow \\
\ldots s_a, s_0^*)(s_{f'(c)},\ s_a^* \ldots \\
\Downarrow \\
\ldots s_a, s_0^*)(s_{f'(c)},\ s_a^* \ldots \\
\Downarrow \\
\ldots s_a, s_0^*)(s_{f'(c)},\ s_a^* \ldots \\
\Downarrow \\
\ldots s_a, s_0^*)(s_0^*,\ s_{f'''(c+1)}^* \ldots \\
\Downarrow \\
\ldots s_a, s_0^*)(s_{c+1}^*,\ s_a^* \ldots
\end{array}
$$

Constructing long polymers

Case:	$b = r,\ c < r$
Insertion sequence:	$s_a,s_r)(s_c,s_a^* \Downarrow_{O(1)} s_a,s_0)(s_{c+1},s_a^*$
Result:	$b = 0,\ {+}{+}c$

This is the easiest case.

A total of 23 monomer type families used.

$$\ldots s_a,s_r)(s_c,s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_r)(s_r^*,s_{f(c)}\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_r)(s_{f'(c)},\ s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_{f''(c)},s_{f'(c)}^*)(s_{f'(c)},\ s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_0^*)(s_{f'(c)},\ s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_0^*)(s_{f'(c)},\ s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_0^*)(s_{f'(c)},\ s_a^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_0^*)(s_0^*,\ s_{f'''(c+1)}^*\ldots$$
$$\Downarrow$$
$$\ldots s_a,s_0^*)(s_{c+1}^*,\ s_a^*\ldots$$

An upper bound for polymer length

- Every insertion sequence in polymer must consist of unique insertion sites **a,b)(c,a*** that each accept a monomer.
- Prove there are $O(k^{3/2})$ such sites.
- Proof idea: maximize |{a}|*|{b}|*|{c}| subject to
    - |{a}|*|{b}|, |{a}|*|{c}| = O(k) since they are monomer right/left halves.
    - |{b}|*|{c}| = O(k) since site needs $(b^*,\_,\_,c^*)^+$ inserted.
- |{a}|*|{b}|*|{c}| = $O(k^{3/2})$

Polymer growth speed

Insertion system:

Monomer types: $(b^*,a^*,a,b)^+$ $(b^*,x,x,b)^+$

Concentrations: 0.5 0.5

Initiator: $(a,b)(b^*,a^*)$

$$(a,b)(b^*,a^*)$$
$$\Downarrow 1$$
$$(a,b)(b^*,a^*,a,b)(b^*,a^*)$$
$$\Downarrow 2$$
$$(a,b)(b^*,a^*,a,b)(b^*,a^*,a,b)(b^*,a^*,a,b)(b^*,a^*)$$
$$\Downarrow 4$$
$$(a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*)$$

Each round of insertions takes O(1) expected time.

Construction of length $n = 2^i-1$ takes O(i) expected time.

O(log(n))

Insertion system:

Monomer types: $(b^*,a^*,a,b)^+$ $(b^*,x,x,b)^+$

Concentrations: 0.5 0.5

Initiator: $(a,b)(b^*,a^*)$

$$(a,b)(b^*,a^*)$$
$$\Downarrow 1$$
$$(a,b)(b^*,a^*,a,b)(b^*,a^*)$$
$$\Downarrow 2$$
$$(a,b)(b^*,a^*,a,b)(b^*,a^*,a,b)(b^*,a^*,a,b)(b^*,a^*)$$
$$\Downarrow$$
$$(a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*,a,b)(b^*,x,x,b)(b^*,a^*)$$

Constructing polymers in $O(\log(n))$ expected time is easy if infinite polymers allowed.

Each round of insertions takes $O(1)$ expected time.

Construction of length $n = 2^i-1$ takes $O(i)$ expected time.

$O(\log(n))$

Constructing polymers quickly

Theorem: a system constructing a finite number of polymers can **deterministically** construct a polymer of length n in:

- $O(\log^3(n))$ expected time [Dabby, Chen 2013]
- $O(\log^{5/3}(n))$ expected time [This work]
- only $\Omega(\log^{5/3}(n))$ expected time [This work]

Constructing polymers quickly

Theorem: a system constructing a finite number of polymers can **deterministically** construct a polymer of length n in:

- $O(\log^3(n))$ expected time [Dabby, Chen 2013]
- $O(\log^{5/3}(n))$ expected time [This work]
- only $\Omega(\log^{5/3}(n))$ expected time [This work]

A lower bound for growth speed

- Consider longest insertion sequence carried out;
  must have length at least $\log_2(n)$.

- By polymer length upper bound, must insert
  $\Omega(\log^{2/3}(n))$ different monomer types.

- Optimization problem: pick #insertions, concentrations of $\log^{2/3}(n)$ monomer types into $\log_2(n)$ sites to minimize total expected time.

- Assuming each site accepts 1 monomer type,
  minimum expected time is $\Omega(\log^{5/3}(n))$.

A lower bound for growth speed

- Consider longest insertion sequence carried out; must have length at least $\log_2(n)$.
- By polymer length upper bound, must insert $\Omega(\log^{2/3}(n))$ different monomer types.
- Optimization problem: pick #insertions, concentrations of $\log^{2/3}(n)$ monomer types into $\log_2(n)$ sites to minimize total expected time.
- Assuming each site accepts 1 monomer type, minimum expected time is $\Omega(\log^{5/3}(n))$.

System is *deterministic.*

Constructing polymers quickly

Theorem: a system constructing a finite number of polymers can **deterministically** construct a polymer of length n in:

- $O(\log^{5/3}(n))$ expected time
- only $\Omega(\log^{5/3}(n))$ expected time

Constructing polymers quickly

Theorem: a system constructing a finite number of polymers can **deterministically** construct a polymer of length n in:

- $O(\log^{5/3}(n))$ expected time

- only $\Omega(\log^{5/3}(n))$ expected time

Theorem: a system constructing a finite number of polymers can **non-deterministically** construct a polymer of length n in:

- $O(\log^{3/2}(n))$ expected time

- only $\Omega(\log^{3/2}(n))$ expected time

Constructing polymers quickly

Theorem: a system constructing a finite number of polymers can **deterministically** construct a polymer of length n in:

- $O(\log^{5/3}(n))$ expected time
- only $\Omega(\log^{5/3}(n))$ expected time

Theorem: a system constructing a finite number of polymers can **non-deterministically** construct a polymer of length n in:

- $O(\log^{3/2}(n))$ expected time
- only $\Omega(\log^{3/2}(n))$ expected time

Speedup by non-determinism

$$\ldots s_1, s_0)(s_0^*, s_1^* \ldots$$

$$\Downarrow_{O(1)}$$

$$\ldots s_1, s_1)(s_1^*, s_1^* \ldots$$

Needs unique monomers.

Unique monomers ⇒ slow.

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3^*, s_6^* \ldots$$

$$\Downarrow_{O(1)}$$

$$\ldots s_6, s_4)(s_4^*, s_6^* \ldots$$

Needs unique monomers.

Unique monomers ⇒ slow.

Speedup by non-determinism

$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3^*, s_6^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}^*, s_6^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3^*, s_6^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}^*, s_6^* \ldots$$

$$\Downarrow \text{ If b = 3, guess a}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}^*, s_6^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ If b = 3, guess a}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ If a = 6}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}{}^*, s_{f(a)}{}^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Big\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Big\Downarrow \text{ If b = 3, guess a}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Big\Downarrow \text{ If a = 6}$$

$$\ldots s_{f(6)}, s_{f(4)})(s_{f(4)}{}^*, s_{f(6)}{}^* \ldots$$

Speedup by non-determinism

$\ldots s_6, s_3)(s_3^*, s_6^* \ldots$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Big\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ If } b \neq 3\text{, insertion sequence done.}$$

$$\times$$

Speedup by non-determinism

$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Big\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

$$\Downarrow \text{ If b = 3, guess a}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}{}^*, s_6{}^* \ldots$$

Speedup by non-determinism

$$\ldots s_6, s_3)(s_3^*, s_6^* \ldots$$

$$\Downarrow \text{ Guess b}$$

$$\ldots s_6, s_3)(s_{f(b+1)}^*, s_6^* \ldots$$

$$\Downarrow \text{ If } b = 3, \text{ guess a}$$

$$\ldots s_{f(a)}, s_{f(b+1)})(s_{f(b+1)}^*, s_6^* \ldots$$

$$\Downarrow \text{ If } a \neq 6, \text{ insertion sequence done.}$$

$$\times$$

Why does guessing give a speed-up?

Speedup by non-determinism

$$\ldots S_6, S_3)(S_3{}^*, S_6{}^* \ldots$$

$$\Downarrow \text{Guess b}$$

$$\ldots S_6, S_3)(S_{f(b+1)}{}^*, S_6{}^* \ldots$$

$$\Downarrow \text{If b = 3, guess a}$$

$$\ldots S_{f(a)}, S_{f(b+1)})(S_{f(b+1)}{}^*, S_6{}^* \ldots$$

$$\Downarrow \text{If a = 6}$$

$$\ldots S_{f(a)}, S_{f(b+1)})(S_{f(b+1)}{}^*, S_{f(a)}{}^* \ldots$$

Speedup by non-determinism

$s_6, s_3)(s_3{}^*, s_6{}^*$

$\Downarrow$ Guess b

$s_6, s_3)(s_{f(b+1)}{}^*, s_6{}^*$

$s_6, s_5)(s_5{}^*, s_6{}^*$

$\Downarrow$ Guess b

$s_6, s_5)(s_{f(b+1)}{}^*, s_6{}^*$

$s_6, s_8)(s_8{}^*, s_6{}^*$

$\Downarrow$ Guess b

$s_6, s_8)(s_{f(b+1)}{}^*, s_6{}^*$

$s_6, s_{11})(s_{11}{}^*, s_6{}^*$

$\Downarrow$ Guess b

$s_6, s_{11})(s_{f(b+1)}{}^*, s_6{}^*$

Speedup by non-determinism

$s_6,s_3)(s_3{}^*,s_6{}^*$ $\quad$ $s_6,s_5)(s_5{}^*,s_6{}^*$

$\Downarrow$ Guess b $\quad$ $\Downarrow$ Guess b

$s_6,s_3)(s_{f(b+1)}{}^*,s_6{}^*$ $\quad$ $s_6,s_5)(s_{f(b+1)}{}^*,s_6{}^*$

The same set of monomers are insertable to *all* of these sites.

$s_6,s_8)(s_8{}^*,s_6{}^*$ $\quad$ $s_6,s_{11})(s_{11}{}^*,s_6{}^*$

$\Downarrow$ Guess b $\quad$ $\Downarrow$ Guess b

$s_6,s_8)(s_{f(b+1)}{}^*,s_6{}^*$ $\quad$ $s_6,s_{11})(s_{f(b+1)}{}^*,s_6{}^*$

Non-determinism speed-up

- Sites $s_a, s_b)(s_b^*, s_a^*$ with $\Theta(k^{1/2})$ values for a, b
    - So $k = \Theta(\log(n))$
- Every site accepts $\Theta(k^{1/2})$ monomers: $\Omega(1/k^{1/2})$ total concentration.
- Expected insertion time is $O(k^{1/2}) = O(\log^{1/2}(n))$.
- Total expected construction time is $O(\log^{3/2}(n))$.

Polymer length vs. growth speed

Types vs Speed Tradeoff

Is it possible to construct a polymer using $O(\log^{2/3}(n))$ monomer types in $O(\log^{3/2}(n))$ expected time?

Types vs Speed Tradeoff

Is it possible to construct a polymer using $O(\log^{2/3}(n))$ monomer types in $O(\log^{3/2}(n))$ expected time?

$O(\log(n))$ types in $O(\log^{3/2}(n))$ expected time.

$O(\log^{2/3}(n))$ types in $O(\log^{5/3}(n))$ expected time.

Types vs Speed Tradeoff

Is it possible to construct a polymer using $O(\log^{2/3}(n))$ monomer types in $O(\log^{3/2}(n))$ expected time?

$O(\log(n))$ types in $O(\log^{3/2}(n))$ expected time.

$O(\log^{2/3}(n))$ types in $O(\log^{5/3}(n))$ expected time.

Theorem: in a system constructing a finite number of polymers, constructing a polymer of length n using k monomer types takes $\Omega(\log^2(n)/k^{1/2})$ expected time.

Types vs Speed Tradeoff

Is it possible to construct a polymer using $O(\log^{2/3}(n))$ monomer types in $O(\log^{3/2}(n))$ expected time?

> $O(\log(n))$ types in $O(\log^{3/2}(n))$ expected time.
>
> $O(\log^{2/3}(n))$ types in $O(\log^{5/3}(n))$ expected time.

<u>Theorem:</u> in a system constructing a finite number of polymers, constructing a polymer of length n using k monomer types takes $\Omega(\log^2(n)/k^{1/2})$ expected time.

> $O(\log^{2/3}(n))$ types $\Rightarrow$ $\Omega(\log^{5/3}(n))$ expected time
>
> $O(\log^{3/2}(n))$ expected time $\Rightarrow$ $\Omega(\log(n))$ types

Conclusions

Conclusions

- Insertion systems and context-free grammars are computationally equivalent.
- Growing finitely takes more monomer types and more time than infinite growth.
- Growing the longest finite polymers possible requires use of novel properties — cannot use CFG equivalence!
- Growing finite polymers fastest needs non-determinism and more monomer types.
- A smooth tradeoff between polymer length and growth rate.

Thank you.

Joint work with Benjamin Hescott and Caleb Malchik.

Some results published in:

C. Malchik, A. Winslow, Tight bounds for active self-assembly with an insertion primitive, Proceedings of 22nd European Symposium on Algorithms, 677-688, 2014.