

# Geometry-Driven Selectivity in Macrocyclization via Rigid-Body Simulation

Yuya Kawamura<sup>1</sup>, Takahiro Iwamoto<sup>2</sup>, and Yohei Nishidate<sup>1,3</sup>[0000–0001–5008–9128]

<sup>1</sup> The University of Aizu, Aizu-Wakamatsu, Fukushima 9658580, Japan

<sup>2</sup> Faculty of Molecular Chemistry and Engineering, Kyoto Institute of Technology, Goshokaido-cho, Matsugasaki, Sakyo-ku, Kyoto 6068585, Japan

<sup>3</sup> Materials Innovation Lab, Kyoto Institute of Technology, Goshokaido-cho, Matsugasaki, Sakyo-ku, Kyoto 6068585, Japan  
[nisidate@u-aizu.ac.jp](mailto:nisidate@u-aizu.ac.jp)

**Abstract.** Macrocyclization under largely irreversible conditions often yields complex mixtures. A recently reported organosilicon system is a striking exception: multiple Si–O bond formations between an aromatic diol and a dichlorosilane produce an almost exclusively square cyclic tetramer at practical concentrations without detectable Si–O exchange. We recast this selectivity problem as geometry-constrained assembly of spherical nodes and cylindrical linkers undergoing stochastic encounters and irreversible bonding. Local preferences for an orthogonal bond angle and coplanarity are encoded either as *rigid* holonomic pose constraints or as *compliant* spring–damper bonds that relax toward the target pose. By varying geometric filters (valency, bond angle, coplanarity), post-bond stabilization (rigid vs. compliant), and density, we find: (i) passive geometric selection is necessary to suppress non-viable aggregates but is insufficient to yield the target in abundance; (ii) replacing rigid joints with compliant spring–damper joints shortens the time to the first target and raises the target-object fraction to  $\sim 7\%$ , while completed cycles remain strongly enriched in squares ( $S_4 \approx 85\%$ ); and (iii) once non-viable bonding pathways are suppressed, increasing density primarily accelerates productive encounters without eroding ring-size selectivity. These results emphasize that extreme selectivity under irreversible conditions emerges from the conjunction of stringent local geometry, post-bond stabilization, and adequate collision frequency.

**Keywords:** Macrocyclization · Rigid-body simulation · Geometric constraints · Spring–damper bonds · Stochastic assembly

## 1 Introduction

### 1.1 Background

Macrocycles are foundational motifs in chemistry, biology, and materials science. They support molecular recognition, supramolecular architectures, catalysis, and

drug discovery. Despite this breadth, methods that deliver a single, well-defined ring size with high yield at practical concentrations remain limited [1, 2]. An organosilicon system recently reported in the literature offers a stringent benchmark for explaining selectivity in macrocyclization. For recent examples, see [3–5].

Iwamoto and co-workers [6] reported that the formation of multiple Si–O bonds between an aromatic diol and a dichlorosilane affords an exclusively square cyclic tetramer with nearly quantitative yield. The reaction proceeds within minutes and remains effective at concentrations up to 3.0 M. Under the reported conditions, the selectivity is not attributable to precipitation or templation, and Si–O bond exchange (metathesis) is inactive. These observations indicate history-dependent assembly rather than equilibration through reversible bonds.

The authors proposed a mechanism in which an Ar–O–Si–O–Ar hemilabile interaction weakly enforces an orthogonal local environment at silicon. Density functional theory and noncovalent interaction analyses support a near-orthogonal minimum (approximately 86°). However, conformational effects alone are insufficient to rationalize the near-quantitative selectivity observed under these largely irreversible and high-concentration conditions [1]. Although several alternatives were ruled out and bond formation appears effectively irreversible, a complete quantitative account of the near-exclusive selectivity remains open and goes beyond what is typically addressed by equilibrated ring-size statistics or local quantum-chemical analyses.

## 1.2 Our contribution

This work reinterprets macrocyclization selectivity as a constrained geometric assembly problem. Spherical and cylindrical building blocks are modeled as rigid bodies with angle and length tolerances motivated by experiment. Event-driven simulations are carried out with stochastic encounters and bond-formation rules. The objective is to determine which geometric and protocol parameters—such as concentration, linker aspect ratio, angular tolerance on bond formation, and capture distance—are sufficient to reproduce the exclusive square tetramer.

Local geometric preferences inferred from quantum chemistry enter the model through a target bond angle  $\mathcal{D}[\theta_0]$  at the sphere. In particular, we choose  $\mathcal{D}[\theta_0] = 90^\circ$  with a finite angular tolerance to represent the nearly orthogonal ( $\approx 86^\circ$ ) Ar–O–Si–O–Ar conformation reported in the experimental and DFT studies [6]. This geometry-first perspective complements chemistry-first explanations by isolating minimal sufficient ingredients and by producing testable predictions about how selectivity should change when these geometric constraints are tightened or relaxed.

Because Si–O bond metathesis is inactive, the system operates far from equilibrium and the assembly is inherently history-dependent. Standard molecular dynamics or density functional theory are well suited to describe local structures and reaction energetics, but they do not directly capture the stochastic, irreversible assembly pathways that determine the final ring-size distribution. Our

rigid-body model is designed as a minimal, computationally accessible framework to probe such history-dependent assembly under simple geometric rules.

## 2 Rigid-body modeling of macrocyclization

### 2.1 Chemical context and coarse-graining rationale

We consider a system in which a rod-like aromatic diol (hydroquinone; 1,4-benzenediol) reacts with a compact organosilicon reagent (dimethyldichlorosilane,  $\text{Me}_2\text{SiCl}_2$ ) to form Si–O linkages (Fig. 1). On the assembly timescale, the aromatic core behaves as a short and stiff segment. It is therefore idealized as a cylindrical linker with two reactive termini. The organosilicon reagent presents a comparatively isotropic steric profile and is idealized as a sphere. This coarse-graining retains dominant steric features (shape, size, and encounter multiplicity) while decoupling atomistic details that are not essential for ring-size selectivity under largely irreversible conditions.

To organize the model, we divide its ingredients into three groups. Geometric conditions  $\mathcal{G}$  specify the shapes, sizes, and boundary conditions of the entities. Initial conditions  $\mathcal{I}$  specify how many entities are present and how they are placed and initialized in velocity space. Dynamic conditions  $\mathcal{D}$  specify how rigid-body collisions are resolved and when bonds are created, including local reactivity constraints such as valency and bond-angle/coplanarity tolerances. Table 1 summarizes this grouping of chemical constituents into coarse-grained entities and their associated  $\mathcal{G}$ ,  $\mathcal{I}$ , and  $\mathcal{D}$  parameters. The remainder of this section introduces  $\mathcal{G}$  (Sec. 2.2),  $\mathcal{I}$  (Sec. 2.3), and  $\mathcal{D}$  (Sec. 2.4) in turn, and then defines the simulation protocols and evaluation metrics used in Sec. 2.5.

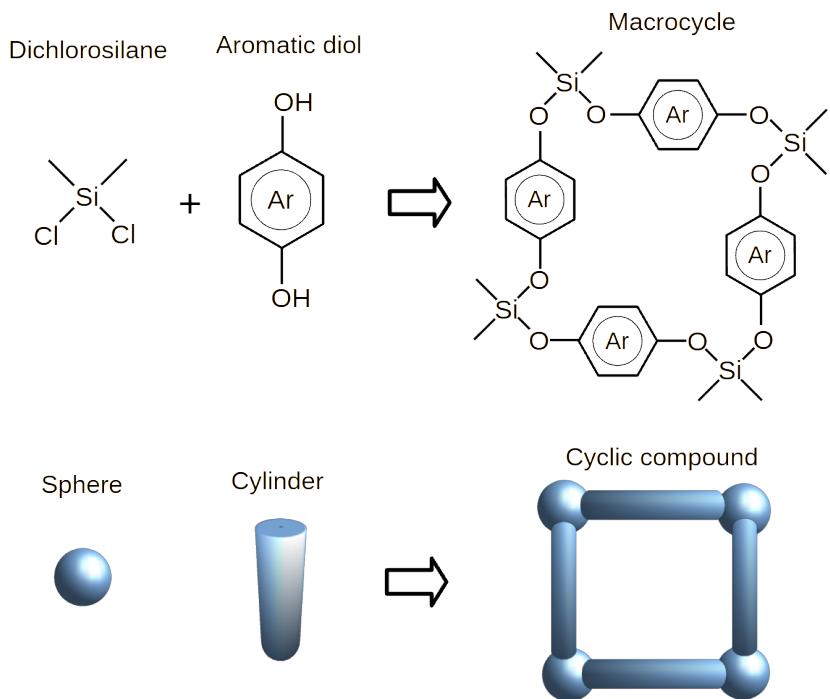
### 2.2 Geometric conditions $\mathcal{G}$

Spheres ( $S$ ) and cylindrical linkers ( $L$ ) are treated as rigid bodies embedded in a box-shaped domain ( $B$ ). The entity geometry is defined by the sphere radius  $\mathcal{G}_S[r]$ , the linker radius  $\mathcal{G}_L[r]$ , and the linker axial length  $\mathcal{G}_L[\ell]$ . The domain is specified by the edge length  $\mathcal{G}_B[\ell]$  and the boundary condition  $\mathcal{G}_B[\beta] \in \{\text{periodic, reflective}\}$  (Table 2).

For contacts, object-wise restitution and friction coefficients are combined into pair values through an operator  $\mathcal{G}[\mathbf{E}]$ :

$$\mathcal{G}_{(X,Y)}[e] := \mathcal{G}[\mathbf{E}](\mathcal{G}_X[e], \mathcal{G}_Y[e]). \quad (1)$$

Standard rigid-body engines offer several options for  $\mathcal{G}[\mathbf{E}]$  (for example, maximum, minimum, product, or arithmetic mean) [7]. Unless stated otherwise, the minimum operator is used for restitution, and the product is used for friction.



**Fig. 1.** Mapping from chemical building blocks to coarse-grained geometric entities. *Top:* a compact organosilicon reagent ( $\text{Me}_2\text{SiCl}_2$ ) and a rod-like aromatic diol (hydroquinone) combine to form a macrocycle through multiple Si–O linkages. *Bottom:* the coarse-grained representation used in this study. The organosilicon reagent is modeled as a sphere ( $S$ ) and the diol as a cylindrical linker ( $L$ ). Valency and bond-angle/coplanarity constraints are imposed at bond formation rather than embedded in the entity shapes.

**Table 1.** Constituents, coarse-grained entities, approximations, and modeling rationale.

Constituent (target)	CG entity (symbol)	Approximation	Rationale
Molecule A (dichlorosilane)	Sphere ( $S$ )	Rigid Isotropic contact Geometric conditions specified by $\mathcal{G}_S[X]$	Compact organosilicon Near-isotropic sterics Vertex constraints at bonding
Molecule B (aromatic diol)	Cylindrical linker ( $L$ )	Rigid, Two termini Geometric conditions specified by $\mathcal{G}_L[X]$	Short, Stiff aromatic core Rod-like sterics dominate
Medium	Domain ( $B$ )	Periodic or Reflective Geometric conditions specified by $\mathcal{G}_B[X]$	Decouple solvent motion Focus on steric/kinematic effects
Dynamic Condition	Event rules (in $\mathcal{D}$ )	Conditions specified by $\mathcal{D}_X[Y_X]$ e.g. valency, bond angle, coplanarity, capture radius	History-dependent assembly
Initial Condition	Initial settings (in $\mathcal{I}$ )	Conditions specified by $\mathcal{I}_X[Y_X]$ e.g. numbers of entities initial positions, angular/translational velocity, hard-overlap rejection	Control composition and preorganization

### 2.3 Initial condition $\mathcal{I}$

We simulate a finite population of  $N_S$  spheres and  $N_L$  cylindrical linkers in a box-shaped domain  $B$ . For a fixed domain size  $\mathcal{G}_B[\ell]$ , the pair  $(N_S, N_L)$  controls both the effective packing fraction (concentration) and the stoichiometric ratio between the two species. More generally, the number density is controlled jointly by the box edge length  $\mathcal{G}_B[\ell]$  (in  $\mathcal{G}$ ) and the particle counts  $(N_S, N_L)$  (in  $\mathcal{I}$ ); in this work we typically vary  $(N_S, N_L)$  at fixed  $\mathcal{G}_B[\ell]$ .

To control initial degrees of mixture, a single mixing parameter  $\mathcal{I}[\alpha] \in [0, 1]$  is employed (Table 3). Let  $(\mathcal{B}_S, \mathcal{B}_L)$  be a fixed equal-volume partition of  $\mathcal{B}$ . For each entity type  $\tau \in \{S, L\}$ , positions are sampled from a mixture of a uniform distribution within  $\mathcal{B}_\tau$  and a uniform distribution within  $\mathcal{B}$ :

$$(1 - \mathcal{I}[\alpha]) P_{\text{seg}}(\tau) + \mathcal{I}[\alpha] P_{\text{uni}},$$

with hard-overlap rejection. Thus,  $\mathcal{I}[\alpha] = 0$  produces full segregation, whereas  $\mathcal{I}[\alpha] = 1$  produces full mixing. At high packing fractions, naive rejection sampling

**Table 2.** Geometry and environment parameters grouped in  $\mathcal{G}$ .

Notation	Meaning	Default
$\mathcal{G}_S[r]$	Sphere radius $r_S$	1.0
$\mathcal{G}_L[r]$	Linker radius $r_L$	1.0
$\mathcal{G}_L[\ell]$	Linker length $\ell_L$	4.0
$\mathcal{G}_B[\ell]$	Box edge length $\ell_B$	50
$\mathcal{G}_B[\beta]$	Boundary condition	reflective
$\mathcal{G}_S[e], \mathcal{G}_L[e], \mathcal{G}_B[e]$	Object-wise restitution	1.0
$\mathcal{G}_S[\mu], \mathcal{G}_L[\mu], \mathcal{G}_B[\mu]$	Object-wise static friction	0.0
$\mathcal{G}_S[\nu], \mathcal{G}_L[\nu], \mathcal{G}_B[\nu]$	Object-wise dynamic friction	0.0
$\mathcal{G}_{(X,Y)}[e]$	Pair restitution via (1)	min
$\mathcal{G}_{(X,Y)}[\mu], \mathcal{G}_{(X,Y)}[\nu]$	Pair friction via (1)	product

**Table 3.** Initialization parameters grouped in  $\mathcal{I}$  (single-parameter mixing).

Notation	Meaning	Default (Range)
$\mathcal{I}_S[N_S], \mathcal{I}_L[N_L]$	Number of spheres/linkers	300 each
$\mathcal{I}[\alpha]$	Mixing degree (0 = segregated, 1 = fully mixed)	1 (0–1)
$\mathcal{I}[\text{init}]$	Velocity initialization method	uniform or thermal
$\mathcal{I}[v_{\max}], \mathcal{I}[\omega_{\max}]$	Limits (for $\mathcal{I}[\text{init}] = \text{uniform}$ )	1.0
$\mathcal{I}[T]$	Temperature (for $\mathcal{I}[\text{init}] = \text{thermal}$ )	1.0 (reduced units)

becomes inefficient because most trial positions overlap existing bodies. In that regime, we use stratified seeding or Poisson-disk sampling to place particles more uniformly and reduce overlaps, while keeping the same intended mixture between segregated and uniform placement (i.e., the same value of  $\mathcal{I}[\alpha]$  and the same routing of samples through  $P_{\text{seg}}$  and  $P_{\text{uni}}$ ).

Translational and angular velocities are sampled independently of positions, with a sampling mode  $\mathcal{I}[\text{init}] \in \{\text{uniform}, \text{thermal}\}$ . In the uniform mode, each Cartesian velocity component and each angular-velocity component is drawn from a symmetric uniform distribution on  $[-\mathcal{I}[v_{\max}], \mathcal{I}[v_{\max}]]$  and  $[-\mathcal{I}[\omega_{\max}], \mathcal{I}[\omega_{\max}]]$ , respectively. In the thermal mode, speeds follow a Maxwell–Boltzmann distribution at temperature  $\mathcal{I}[T]$  and directions are chosen isotropically, following standard practice in molecular dynamics simulations [9–12]. Unless otherwise stated, we use the uniform mode as a neutral baseline that avoids committing to a particular effective temperature. For reproducibility, the realized mixing is reported as the fraction of cross-type nearest-neighbor pairs in the initial contact graph.

## 2.4 Dynamics $\mathcal{D}$

Event-driven rigid-body dynamics are used, with ballistic motion between events (Table 4). We denote by  $\theta$  the local bond angle at a sphere, i.e., the angle between the axes of two linker termini bonded to the same sphere, and by  $\phi$  the dihedral angle that measures deviation from coplanarity of the four-point sequence L–S–L–S.

A candidate bond between a linker end and a sphere is detected when the closest approach falls below the capture distance  $\mathcal{D}_{(L,S)}[r_{c(L,S)}]$ . Bond formation requires valency satisfaction and two geometric conditions: the bond angle  $\theta$  must lie within  $\mathcal{D}[\theta_0] \pm \mathcal{D}[\Delta_\theta]$ , and the deviation from coplanarity must lie within  $\pm \mathcal{D}[\Delta_\phi]$ . Once created, bonds are irreversible unless explicitly relaxed. Upon bond creation, incompatible predicted events are canceled and the bond graph is updated; simple cycles are tracked for downstream metrics. This irreversibility, together with finite bond-angle and coplanarity tolerances, produces a history-dependent assembly process that mirrors the experimental absence of Si–O bond exchange.

In addition to these passive geometric constraints, the model optionally supports active constraint forces implemented as linear and angular spring–damper terms on the relative pose of bonded pairs. Two global parameter pairs,  $(\mathcal{D}[k], \mathcal{D}[c])$  and  $(\mathcal{D}[k_\omega], \mathcal{D}[c_\omega])$ , control the stiffness and damping of translational and rotational corrections, respectively. When  $\mathcal{D}[k] > 0$  and/or  $\mathcal{D}[k_\omega] > 0$ , a newly formed bond is associated with a virtual constraint that penalizes deviations of the relative position and orientation of the bonded bodies from a prescribed target pose. This target pose is chosen such that the line connecting the linker termini and the sphere center, and the local frame of the linker pair, realize the target bond angle  $\mathcal{D}[\theta_0]$  and a coplanar arrangement ( $\phi \approx 0$ ). Translational spring–damper forces, governed by  $(\mathcal{D}[k], \mathcal{D}[c])$ , oppose displacements of the bond anchor points, while rotational spring–damper torques, governed by  $(\mathcal{D}[k_\omega], \mathcal{D}[c_\omega])$ , oppose rotations away from the target orientation. We use two bonding regimes: (i) a *rigid* regime, in which bonds are enforced as holonomic pose constraints without elastic compliance, and (ii) a *compliant* regime, in which the relative pose is penalized by spring–damper terms with finite  $(\mathcal{D}[k], \mathcal{D}[c], \mathcal{D}[k_\omega], \mathcal{D}[c_\omega])$  that relax toward the target angle and coplanarity. For clarity, “rigid” means no post-bond elastic relaxation; “compliant” means small deviations are allowed and dissipatively driven back toward  $\mathcal{D}[\theta_0]$  and coplanarity.

## 2.5 Protocol and evaluation metrics

Experimental conditions are specified by  $(\mathcal{G}, \mathcal{I}, \mathcal{D})$ . We vary the concentration and the stoichiometric ratio (through  $N_S$  and  $N_L$ ), the linker aspect ratio  $\mathcal{G}_L[\ell]/(2\mathcal{G}_L[r])$ , the geometric tolerances  $(\mathcal{D}[\Delta_\theta], \mathcal{D}[\Delta_\phi])$ , the capture distance  $\mathcal{D}_{(L,S)}[r_{c(L,S)}]$ , and the initialization options. Each setting is replicated  $R$  times with independent seeds.

**Table 4.** Dynamic parameters grouped in  $\mathcal{D}$  (bond-formation rules, bond-angle/coplanarity tolerances, and optional active constraint forces).

Notation	Meaning	Default
$\mathcal{D}_{(L,S)}[r_{c(L,S)}]$	Capture distance	0.1
$\mathcal{D}_S[v_{as}]$	Sphere valency	2
$\mathcal{D}_L[val]$ , $\mathcal{D}_L[v_{bL}]$	Linker valency at each end	1 each
$\mathcal{D}[\theta_0]$	Target bond angle	90°
$\mathcal{D}[\Delta_\theta]$	Bond-angle tolerance (half-width)	5°
$\mathcal{D}[\Delta_\phi]$	Coplanarity tolerance (half-width)	5°
$\mathcal{D}[k]$	Translational spring constant	0
$\mathcal{D}[c]$	Translational damping coefficient	0
$\mathcal{D}[k_\omega]$	Rotational spring constant	0
$\mathcal{D}[c_\omega]$	Rotational damping coefficient	0

Let  $G_t = (V_t, E_t)$  denote the bond graph at time  $t$ , where vertices represent bodies ( $S$  or  $L$ ) and edges represent bonds. An *object* is a connected component of  $G_t$ . The target object is a single cycle with alternating species that contains four spheres and four linkers (i.e.,  $|S| = |L| = 4$ ). The target-object fraction  $\Phi_{\text{obj}}$  at time  $t$  is

$$\Phi_{\text{obj}}(t) = \frac{N_{\text{tar}}(t)}{N_t}, \quad (2)$$

where  $N_t$  is the total number of connected components and  $N_{\text{tar}}(t)$  is the number of components that satisfy the target definition.

A component is classified as non-viable if any of the following conditions holds.

- NV1:** The maximum degree within the component exceeds two (branching).
- NV2:** The component contains a simple cycle whose alternating length is not eight (incorrect ring size).
- NV3:** The counts of spheres or linkers exceed four (species imbalance).

Otherwise, the component is viable. The viable-object fraction  $\Phi_{\text{viab}}$  at time  $t$  is

$$\Phi_{\text{viab}}(t) = \frac{N_{\text{viab}}(t)}{N_t}. \quad (3)$$

Let  $N_m$  denote the number of completed cycles that contain  $m$  spheres at the final observation time, and let  $N_{\text{cyc}} = \sum_{i \geq 3} N_i$  be the total number of completed cycles with at least three spheres. Then, the selectivity for the square macrocycle is defined as

$$S_4 = \frac{N_4}{N_{\text{cyc}}}. \quad (4)$$

If no cycles are observed in a given condition (i.e.,  $N_{\text{cyc}} = 0$ ), the quantity  $S_4$  is not defined.

Additional observables include the time to first target  $\tau_{\text{tar}}$  and the residual monomer fraction  $M_{\text{mono}}(t)$ . For each replicate  $r$ , the time to first target is defined as

$$\tau_{\text{tar}}^{(r)} = \inf\{ t : N_{\text{tar}}^{(r)}(t) \geq 1 \},$$

i.e., the earliest simulation time at which at least one target object appears. If no target object is formed before the simulation terminates at  $t_{\text{end}}$ , we treat  $\tau_{\text{tar}}^{(r)}$  as right-censored and set  $\tau_{\text{tar}}^{(r)} = t_{\text{end}}$  for the purpose of summarization, indicating in tables that  $\tau_{\text{tar}}$  exceeds  $t_{\text{end}}$ .

Let  $N_{\text{mono},S}(t)$  and  $N_{\text{mono},L}(t)$  denote the numbers of spheres and linkers, respectively, that are monomeric at time  $t$  (degree zero in the bond graph). The residual monomer fraction is defined as

$$M_{\text{mono}}(t) = \frac{N_{\text{mono},S}(t) + N_{\text{mono},L}(t)}{N_S + N_L},$$

which measures the fraction of all entities that remain unbonded at time  $t$ . Species-resolved versions,  $N_{\text{mono},S}(t)/N_S$  and  $N_{\text{mono},L}(t)/N_L$ , can also be inspected when needed, but for brevity we focus on the combined fraction  $M_{\text{mono}}(t)$  and typically report its value at the final observation time  $t_{\text{end}}$ .

In addition to these object-level observables, one could in principle analyze the diversity of assembly pathways via information-theoretic measures such as pathway entropy and kinetic funneling indices; we outline such trajectory-level descriptors as future work in Sec. 4.

### 3 Experiments

We implemented our rigid-body solver using NVIDIA PhysX [8]. Scenario 1 serves as our baseline configuration and imposes only the minimal rules necessary for assembly, and Scenarios 2–6 then introduce additional constraints and protocol variations to determine when selectivity emerges and how sensitive it is to geometric constraints, post-bond stabilization, and density. Unless otherwise noted, default parameter settings are used from Tables 2–4: the numbers of spheres and linkers ( $N_S = N_L = 300$  and thus a 1:1 stoichiometric ratio), fully mixed initialization ( $\mathcal{I}[\alpha] = 1.0$ ), reflective boundary conditions ( $\mathcal{G}_B[\beta] = \text{reflective}$ ), linker aspect ratio  $\mathcal{G}_L[\ell]/(2\mathcal{G}_L[r]) = 2.0$ , capture distance  $\mathcal{D}_{(L,S)}[r_{c(L,S)}] = 0.1$ , bond-angle tolerance  $\mathcal{D}[\Delta_\theta] = 5^\circ$ , and coplanarity tolerance  $\mathcal{D}[\Delta_\phi] = 5^\circ$ . Unless otherwise noted, bonds are modeled as *rigid constraints* with no post-formation compliance. Where stated, we also consider a *compliant* variant in which bonds are spring-damper joints of finite stiffness and damping.

Each experimental condition is replicated  $R = 1000$  times with independent random seeds. Simulations terminate when steady state is detected via the marginal standard error rule (MSER) [13] applied to the total number of connected components, or after a maximum of  $5 \times 10^5$  time steps. Performance is quantified by the target-object fraction  $\Phi_{\text{obj}}$ , selectivity  $S_4$ , viable-object fraction  $\Phi_{\text{viab}}$ , residual monomer fraction  $M_{\text{mono}}$ , the total number of connected components  $N_t$ , and the time to first target  $\tau_{\text{tar}}$ , as defined in Sec. 2.5. In Table 5,

we report all metrics as mean  $\pm$  95% confidence intervals over the  $R = 1000$  replicates.

To systematically investigate the origins of macrocyclization selectivity, we organize our study into a small set of scenarios that differ only in their dynamic constraints and protocol-level parameters. Scenario 1 serves as the baseline with valency-only control. Scenarios 2–4 add passive bond-angle and coplanarity constraints in a stepwise fashion. Scenario 5 augments the most selective passive scenario by switching the bonding rule from rigid to compliant spring-damper joints. Scenario 6 explores density dependence under the same passive geometric selection. Table 5 compiles, for all scenarios, the parameter deviations from the defaults in Tables 2–4 together with the values of the metrics defined in Sec. 2.5. In the scenario-specific discussion below, we refer to Table 5 when comparing outcomes across different constraint sets and densities.

*Scenario 1 (baseline: valency constraint only).* In the baseline model, bond formation is controlled solely by the capture distance and valency constraints. A candidate bond between a linker end and a sphere forms irreversibly when the closest approach falls below  $D_{(L,S)}[r_{c(L,S)}]$  and both the sphere and the linker end have remaining valency ( $D_S[v_{aS}] = 2$ ,  $D_L[v_{aL}] = D_L[v_{bL}] = 1$ ). The local bond angle  $\theta$  at the sphere and the dihedral  $\phi$  are left unconstrained; no explicit angular or coplanarity constraints are applied. This scenario defines a topological baseline that already suppresses unlimited aggregation but does not specifically favor any ring size.

Quantitative results for Scenario 1 are reported in the first row of Table 5. A representative steady-state configuration is shown in the first (top left) panel of Fig. 2. The system is dominated by disordered chains and branched networks. The residual monomer fraction is essentially zero, indicating that almost all entities participate in some aggregate, but only a modest fraction of components are viable ( $\Phi_{\text{viab}} \approx 26\%$ ). Target macrocycles do appear, yet they account for only a few percent of all connected components ( $\Phi_{\text{obj}} \approx 6\%$ ), and only about one quarter of completed cycles are square ( $S_4 \approx 24\%$ ). Thus, valency control alone allows some macrocyclization but with low selectivity and many off-target structures.

*Scenario 2 (baseline + bond-angle constraint).* Relative to Scenario 1, we impose a local bond-angle constraint motivated by the nearly orthogonal Ar–O–Si–O–Ar geometry. For bonds beyond the first at a given sphere, the inter-linker bond angle must satisfy

$$|\theta - D[\theta_0]| \leq D[\Delta_\theta]$$

with  $D[\theta_0] = 90^\circ$  and  $D[\Delta_\theta] = 5^\circ$ . Coplanarity remains unconstrained and active forces are disabled. Under this condition, orthogonality at each bonding site is enforced, but the growing object can freely rotate in three dimensions.

Results for Scenario 2 are summarized in the second row of Table 5, and a representative steady-state configuration is shown in the second (top right) panel of Fig. 2. Compared to Scenario 1, the number of connected components at

steady state increases and the fraction of viable objects nearly doubles ( $\Phi_{\text{viab}} \approx 48\%$ ). The target-object fraction  $\Phi_{\text{obj}}$  remains on the order of a few percent, but the ring-size distribution becomes more focused: squares account for almost half of all completed cycles ( $S_4 \approx 47\%$ ). Thus, a local bond-angle filter strongly reduces non-productive configurations and improves ring-size selectivity, but it does not by itself yield a high fraction of target macrocycles among all objects.

*Scenario 3 (baseline + coplanarity constraint).* Scenario 3 instead imposes only a coplanarity constraint on top of Scenario 1. The dihedral deviation must satisfy

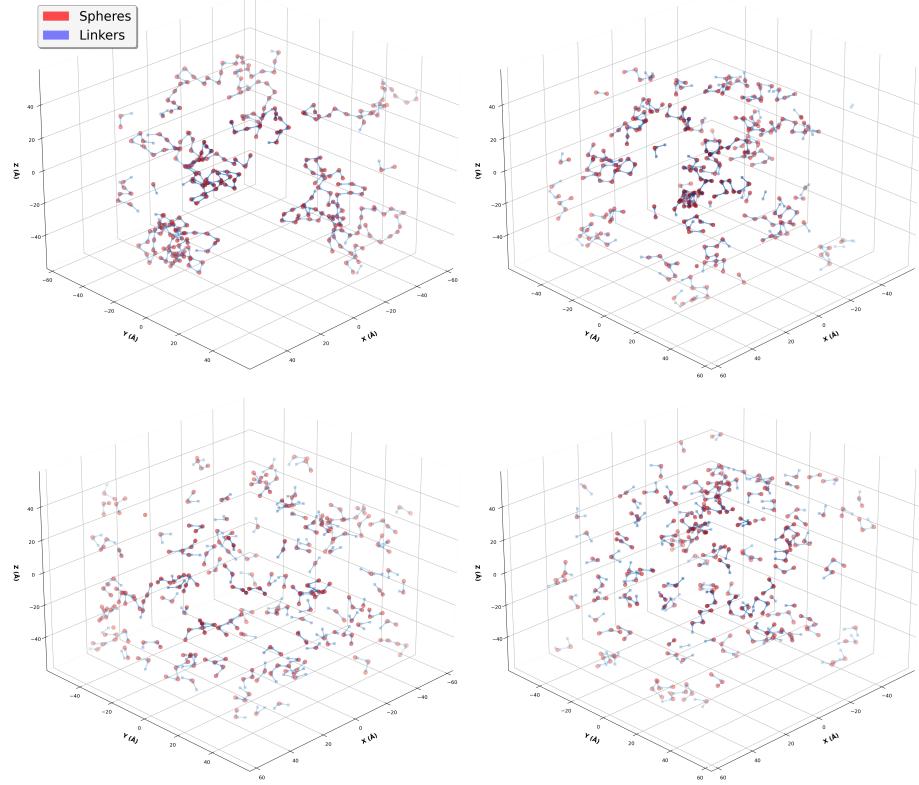
$$|\phi| \leq \mathcal{D}[\Delta_\phi]$$

with  $\mathcal{D}[\Delta_\phi] = 5^\circ$ , while the local bond angle  $\theta$  remains unrestricted. Active forces are disabled. This scenario isolates the effect of dimensionality reduction by confining the assembly to a quasi-two-dimensional plane while leaving branching angles effectively random.

Quantitative results for Scenario 3 are reported in the third row of Table 5, and a representative steady-state configuration is shown in the third (bottom left) panel of Fig. 2. The number of connected components  $N_t$  increases substantially relative to Scenarios 1 and 2, and the viable-object fraction becomes very high ( $\Phi_{\text{viab}} \approx 94\%$ ), indicating that most aggregates respect the degree and size constraints. At the same time, the absolute fraction of target objects is very small ( $\Phi_{\text{obj}} \lesssim 1\%$ ). Among the few completed cycles, squares are favored ( $S_4 \approx 57\%$ ), but ring-closure events are rare compared to the total population of planar oligomers. Coplanarity alone therefore promotes viable, sheet-like architectures while leaving macrocycle formation kinetically suppressed.

*Scenario 4 (complete passive geometric selection).* Scenario 4 combines both bond-angle and coplanarity constraints on top of the valency restrictions in Scenario 1. Bonds form only when the capture distance, valency, bond-angle, and coplanarity conditions are simultaneously satisfied, and bonds are treated as rigid constraints with no compliance. This scenario most closely mirrors the chemically inferred situation in which specific local geometries are enforced through steric and hemilabile interactions around silicon, but with purely passive constraints acting at bond formation.

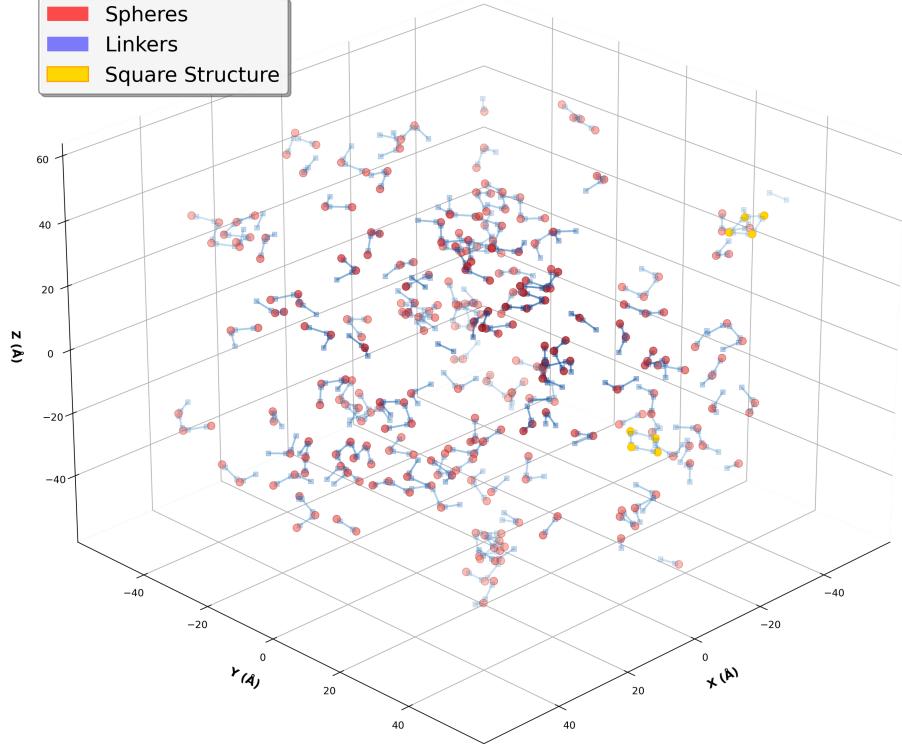
Results for Scenario 4 are listed in the fourth row of Table 5, and a representative steady-state configuration is shown in the fourth (bottom right) panel of Fig. 2. The viable-object fraction approaches unity ( $\Phi_{\text{viab}} \approx 97\%$ ), and the system consists mainly of short, angular, planar oligomers that resemble open fragments of the square macrocycle. The time to first target  $\tau_{\text{tar}}$  increases relative to Scenarios 1–3, consistent with stricter geometric filtering. Although the target-object fraction remains below one percent, the cycles that do form are predominantly square ( $S_4 \gtrsim 80\%$ ). Complete passive geometric selection is therefore very effective at suppressing non-viable architectures and steering rings toward the desired size, but it still yields a small absolute abundance of target macrocycles.



**Fig. 2.** Representative steady-state configurations for Scenarios 1–4. From left to right: Scenario 1 (baseline, valency only), Scenario 2 (baseline + bond-angle constraint), Scenario 3 (baseline + coplanarity constraint), and Scenario 4 (complete passive geometric selection). Spheres ( $S$ ) and cylindrical linkers ( $L$ ) are colored by species, and bonds are shown as rigid connections.

*Scenario 5 (complete selection + compliant bonds).* We now switch from rigid bonds to *compliant* spring-damper bonds of finite stiffness and damping, allowing small post-bond deviations that relax toward the target angle and coplanarity. We activate translational and rotational spring-damper corrections with  $(\mathcal{D}[k], \mathcal{D}[c]) = (1000, 100)$  and  $(\mathcal{D}[k_\omega], \mathcal{D}[c_\omega]) = (1000, 100)$ . For each new bond, the target relative orientation is chosen to realize the preferred bond angle  $\mathcal{D}[\theta_0]$  and coplanar alignment, so that the rotational spring-damper terms oppose deviations of the local bond angle and dihedral from their prescribed values. This represents a geometry-stabilized assembly protocol, in which both bond formation and post-formation relaxation favor the square macrocycle.

Quantitative results for Scenario 5 are given in the fifth row of Table 5. A representative steady-state configuration is shown in Fig. 3. Relative to the rigid-bond baseline, the time to first target is reduced by approximately a factor of



**Fig. 3.** Representative steady-state configuration of Scenario 5.

four, and the fraction of target objects increases almost by an order of magnitude (from well below 1% to about 7%). At the same time, the viable-object fraction remains above 90%, and the ring-size distribution becomes even more focused on squares ( $S_4 \approx 85\%$ ). These results indicate that once passive constraints have eliminated most non-viable pathways, compliant bonds that stabilize the preferred local geometry can substantially increase both the rate and the yield of macrocycle formation without sacrificing selectivity.

*Scenario 6 (complete selection + density sweep).* Finally, Scenario 6 explores density dependence under the same passive geometric selection as Scenario 4 (i.e., bonds treated as rigid constraints). The total number of particles  $N = N_S + N_L$  is varied over  $\{300, 600, 1200\}$ , corresponding to  $(N_S, N_L) = (150, 150)$ ,  $(300, 300)$ , and  $(600, 600)$  at fixed box size  $G_B[\ell] = 50.0$ , while maintaining an approximately stoichiometric ratio  $N_S : N_L \simeq 1 : 1$ . Classically, the Ruggli–Ziegler principle suggests that high dilution favors macrocyclization over polymerization [1]. In the present setting, strict geometric constraints strongly suppress non-viable

bonding, so changing the density primarily affects encounter frequencies and thus the kinetics of ring closure.

Quantitative results for Scenario 6 are reported in the last three rows of Table 5. The case  $(N_S, N_L) = (300, 300)$  coincides with Scenario 4 and is not repeated in the table body. At the lowest density ( $N_S = N_L = 150$ ), aggregates are almost exclusively viable ( $\Phi_{\text{viab}} \approx 100\%$ ), but target macrocycles are very rare ( $\Phi_{\text{obj}} \approx 0.25\%$ ), and the time to first target is comparatively long. When rings do form, however, squares consistently dominate the ring-size distribution with a tight confidence interval ( $S_4 = 83.8\% \pm 1.04\%$ ). Increasing the density to  $(N_S = N_L = 600)$  roughly doubles the number of connected components and shortens the time to first target, while keeping the viable-object fraction above 94%. The fraction of target objects increases compared to the baseline density, and the square selectivity  $S_4$  remains above 80% with a narrower confidence interval. Thus, once non-viable bonding is strongly suppressed by passive geometric selection, higher density facilitates macrocyclization by increasing productive collision frequencies, without leading to uncontrolled polymerization or a loss of ring-size selectivity.

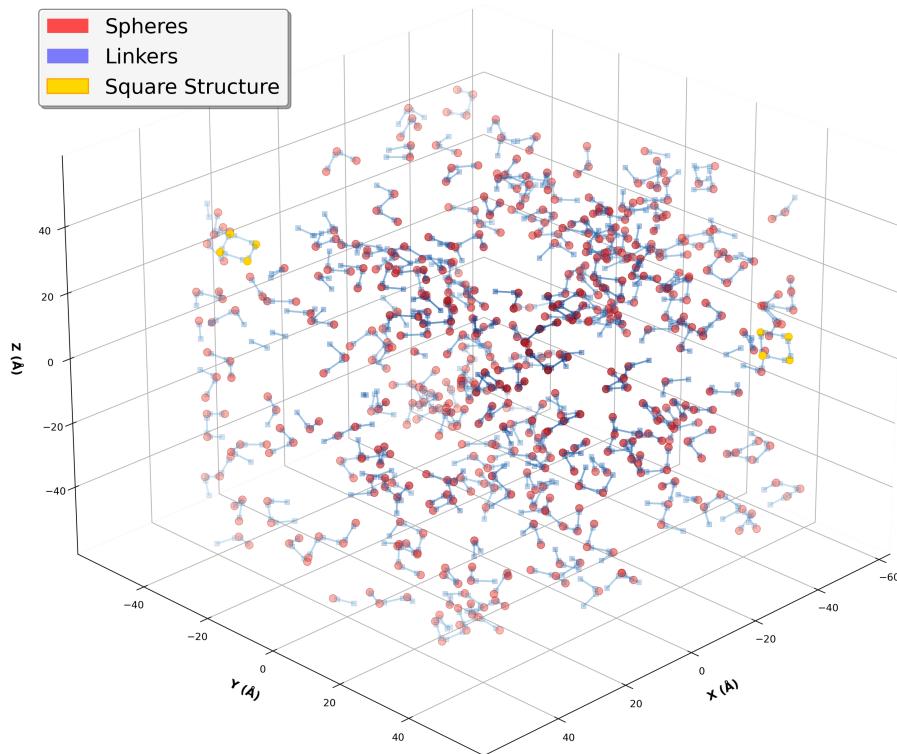
## 4 Conclusion

We examined geometry-driven selectivity in macrocyclization with a minimal rigid-body model that separates geometry, initialization, and dynamics. The model uses spherical nodes and cylindrical linkers, irreversible bond formation, and local rules for valency, bond angle, and coplanarity. This separation allowed us to quantify how each ingredient influences the appearance of a single ring size.

Across scenarios, passive geometric filters are necessary but not sufficient. Valency control alone yields efficient monomer consumption but many disordered aggregates. Adding either a bond-angle constraint or a coplanarity constraint increases the fraction of viable objects yet leaves the absolute abundance of target squares very low. Enforcing both constraints produces short, planar oligomers that match the local square geometry, and cycles that do form are mostly squares, but the overall target-object fraction remains below one percent.

Active stabilization changes the outcome. Replacing rigid joints with compliant spring-damper joints of finite stiffness and damping shortens the time to the first target and increases the target-object fraction to about seven percent, while completed cycles remain strongly enriched in squares ( $S_4 \approx 85\%$ ). A density sweep under passive selection shows that once non-viable bonding is strongly suppressed, increasing encounter frequency helps ring closure without loss of selectivity: at low density targets are rare but square-dominant ( $S_4 = 83.8\%$ ), and at higher density targets appear more often while the viable-object fraction stays high.

These results indicate that selective square macrocyclization under irreversible conditions requires three elements working together: strict local geometry, stabilization of the preferred pose after bonding, and a density high enough to



**Fig. 4.** Representative steady-state configuration of Scenario 6.

overcome kinetic limitations. The framework is intentionally minimal and can be extended to map coarse-grained parameters to experimental observables, to multi-component or templated architectures, and to trajectory-level analyses that probe pathway diversity and kinetic funneling.

## References

- Marti-Centelles, V., Pandey, M. D., Burguete, M. I., Luis, S. V.: Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization. *Chemical Reviews* **115**, 8736–8834 (2015) <https://doi.org/10.1021/acs.chemrev.5b00056>
- Garnes-Portolés, F., Leyva-Pérez, A.: Macrocyclization Reactions at High Concentration ( $\geq 0.2\text{M}$ ): The Role of Catalysis. *ACS Catalysis* **13**(14), 9415–9426 (2023) <https://doi.org/10.1021/acscatal.3c02006>
- Zhao, G., Chen, S., Zhao, W., Li, B., Zhang, W., Zheng, B., Yang, X., Wu, B.: Modular Synthesis of Tetraurea and Octaurea Macrocycles Encoded with Specific Monomer Sequences. *CCS Chemistry* **4**(7), 2498–2507 (2022) <https://doi.org/10.31635/ccschem.021.202101131>

4. Yang, Y., Ying, H., Li, Z., Wang, J., Chen, Y., Luo, B., Gray, D.L., Ferguson, A., Chen, Q., Cheng, J.: Near quantitative synthesis of urea macrocycles enabled by bulky N-substituent. *Nature Communications* **12**(1), 1572 (2021) <https://doi.org/10.1038/s41467-021-21678-3>
5. Yuan, L., Feng, W., Yamato, K., Sanford, A., Xu, D., Guo, H., Gong, B., Highly efficient, one-step macrocyclizations assisted by the folding and preorganization of precursor oligomers. *J. Am. Chem. Soc.* **126**(36), 11120 (2004) <https://doi.org/10.1021/ja0474547>
6. Iwamoto, T., Amano, S., Maeda, K., Shibama, N., Sekiguchi, W., Kazama, Y., Nakamura, Y., Sugamata, K., Imoto, H., Naka, K., Ishii, Y.: Exclusive macrocyclization through multiple Si–O bond formations from diol and dichlorosilane. *Chemical Communications* **61**, 8180–8183 (2025)
7. Unity Documentation, “How collider surface values combine,” <https://docs.unity3d.com/Manual/collider-surfaces-combine.html>, last accessed 2025/10/31
8. NVIDIA PhysX, <https://github.com/NVIDIA-Omniverse/PhysX/>, last accessed 2025/10/31
9. Allen, M.P., Tildesley, D.J.: Computer Simulation of Liquids. 2nd edn. Oxford University Press, Oxford (2017)
10. Frenkel, D., Smit, B.: Understanding Molecular Simulation: From Algorithms to Applications. 2nd edn. Academic Press, San Diego (2002)
11. Shell, M.S.: Molecular dynamics lecture notes. University of California, Santa Barbara, [https://sites.engineering.ucsb.edu/~shell/che210d/Molecular\\_dynamics.pdf](https://sites.engineering.ucsb.edu/~shell/che210d/Molecular_dynamics.pdf), last accessed 2025/10/31
12. Hjorth Larsen, A. et al.: Molecular dynamics tutorial in ASE documentation, [https://ase-lib.org/examples\\_generated/tutorials/md.html](https://ase-lib.org/examples_generated/tutorials/md.html), last accessed 2025/10/31
13. Chodera, J.D.: A Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **12**, 1799–1805 (2016) <https://doi.org/10.1021/acs.jctc.5b00784>

**Table 5.** Representative scenarios and outcomes in a single compact view. Deviations are listed against the defaults in Tables 2–4. Metrics are reported as mean  $\pm$  95% confidence interval over  $R = 1000$  replicates.

Scenario	Deviations from defaults	Key metrics (Sec. 2.5)
1: Baseline (valency restriction only)	No bond-angle or coplanarity constraints or spring-damper constants	$N_t = 31.5688 \pm 0.3635$ $\tau_{\text{tar}} = (0.51 \pm 0.13) \times 10^3$ $\Phi_{\text{obj}} = (5.86 \pm 0.29)\%$ $S_4 = (24.4 \pm 1.14)\%$ $\Phi_{\text{viab}} = (26.3 \pm 0.54)\%$ $M_{\text{mono}} = (0.05 \pm 0.01)\%$
2: Bond-angle constraint	Scenario 1 plus bond-angle constraint: $\mathcal{D}[\theta_0] = 90^\circ$ , $\mathcal{D}[\Delta_\theta] = 5^\circ$	$N_t = 52.9429 \pm 0.2843$ $\tau_{\text{tar}} = (1.08 \pm 0.18) \times 10^3$ $\Phi_{\text{obj}} = (5.69 \pm 0.21)\%$ $S_4 = (47.45 \pm 1.42)\%$ $\Phi_{\text{viab}} = (47.93 \pm 0.39)\%$ $M_{\text{mono}} = (0.14 \pm 0.01)\%$
3: Coplanarity constraint	Scenario 1 plus coplanarity constraint: $\mathcal{D}[\Delta_\phi] = 5^\circ$	$N_t = 120.4247 \pm 0.4125$ $\tau_{\text{tar}} = (1.35 \pm 0.22) \times 10^3$ $\Phi_{\text{obj}} = (0.55 \pm 0.05)\%$ $S_4 = (56.68 \pm 3.5)\%$ $\Phi_{\text{viab}} = (93.86 \pm 0.17)\%$ $M_{\text{mono}} = (0.21 \pm 0.02)\%$
4: Bond-angle and Coplanarity constraints	Scenario 1 plus bond-angle and coplanarity constraints $\mathcal{D}[\theta_0] = 90^\circ$ , $\mathcal{D}[\Delta_\theta] = 5^\circ$ , and $\mathcal{D}[\Delta_\phi] = 5^\circ$	$N_t = 134.0175 \pm 0.4549$ $\tau_{\text{tar}} = (2.34 \pm 0.38) \times 10^3$ $\Phi_{\text{obj}} = (0.37 \pm 0.04)\%$ $S_4 = (80.34 \pm 3.64)\%$ $\Phi_{\text{viab}} = (97.27 \pm 0.11)\%$ $M_{\text{mono}} = (0.35 \pm 0.02)\%$
5: Selection + active forces	Scenario 4 plus compliant bonds: $\mathcal{D}[k] = \mathcal{D}[k_\omega] = 1000$ , $\mathcal{D}[c] = \mathcal{D}[c_\omega] = 100$	$N_t = 112.7550 \pm 0.4112$ $\tau_{\text{tar}} = (0.55 \pm 0.06) \times 10^3$ $\Phi_{\text{obj}} = (6.97 \pm 0.17)\%$ $S_4 = (84.93 \pm 0.86)\%$ $\Phi_{\text{viab}} = (91.57 \pm 0.19)\%$ $M_{\text{mono}} = (0.85 \pm 0.03)\%$
6-1: Selection + density sweep	Scenario 4 with $N_S = N_L = 150$ : fixed $\mathcal{G}_B[\ell] = 50.0$	$N_t = 76.7778 \pm 0.7387$ $\tau_{\text{tar}} = (3.28 \pm 0.69) \times 10^3$ $\Phi_{\text{obj}} = (0.25 \pm 0.18)\%$ $S_4 = (83.8 \pm 1.04)\%$ $\Phi_{\text{viab}} = (99.7 \pm 0.2)\%$ $M_{\text{mono}} = (1.68 \pm 0.24)\%$
6-2: Selection + density sweep	Scenario 4 with $N_S = N_L = 300$ : fixed $\mathcal{G}_B[\ell] = 50.0$	See Scenario 4 above (Default density)
6-3: Selection + density sweep	Scenario 4 with $N_S = N_L = 600$ : fixed $\mathcal{G}_B[\ell] = 50.0$	$N_t = 240.3333 \pm 0.6120$ $\tau_{\text{tar}} = (2.04 \pm 0.57) \times 10^3$ $\Phi_{\text{obj}} = (0.61 \pm 0.37)\%$ $S_4 = (82.5 \pm 2.16)\%$ $\Phi_{\text{viab}} = (94.85 \pm 0.91)\%$ $M_{\text{mono}} = (0.06 \pm 0.03)\%$