MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **96** (2026) 55–74
ISSN: 0340–6253
doi: 10.46793/match.96-1.16525

Semi-Double Trace and Double Cover: A Strong Stable Model for Multi-Component Polypeptide Cages

Saihua Liu[a], **Jiajun Ning**[b,*], **Xiao-Sheng Cheng**[c]

[a] *School of Mathematics and Computational Science, Wuyi University, Jiangmen, Guangdong, 529020, P. R. China*

[b] *School of Mathematical Sciences, Xiamen University, Xiamen, Fujian, 361005, P. R. China*

[c] *School of Mathematics and Statistics, Huizhou University, Huizhou, Guangdong, 516007, P. R. China*

lsh1808@163.com, jiajunning@126.com, chengxs@hzu.edu.cn

(Received July 7, 2025)

Abstract

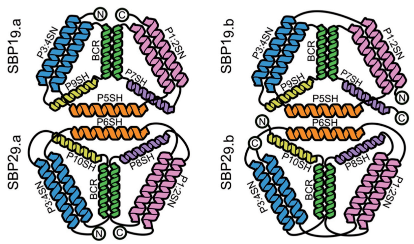
We introduce semi-double trace and semi-strong trace inspired by newly synthesized coiled-coil protein cages, a graph-theoretic generalization of double trace, to model polypeptide nanostructure self-assembly. Through multi-component double covers built from semi-strong traces, we establish a topological assembly framework. Crucially, we prove that double covers attain strong stability when all components are semi-strong. This model resolves a fundamental design challenge: determining oligomeric states via component counts while establishing stability criteria through cyclic vertex figures. These insights provide rigorous principles for engineering biomimetic nanomaterials with programmable topological stability.

*Corresponding author.

1 Introduction

The self-assembly of biopolymers into nanostructures represents a frontier in biomolecular engineering. While DNA nanotechnology has established programmable geometric designs [1, 2, 6, 8, 11, 17, 18], seminal work by Gradišar et al. demonstrated this potential through a tetrahedral cage [7], where twelve helical-binding segments traverse triangular face edges twice. Klavžar employed stable trace to develop a mathematical model that mimics this self-assembly approach for nanostructure design [9]. This model has been refined to better represent these structures [3–5, 13, 14].

Subsequent advances by Lapenta et al. [10] extended this paradigm with a triangular bipyramid assembled from preorganized coiled-coil modules (See Figure 1). Their decomposition into asymmetric or pseudosymmetric subunits revealed the critical role of conformational flexibility: interfacial positioning at N/C-termini preserved structural integrity, whereas rigid topologies led to assembly failure.



(*a*) (*b*)

Figure 1. (*a*) $SBP1_{9.a}$ and $SBP2_{9.a}$ use a rigid closed-loop design: N/C-termini anchored at non-interface vertices restrict interfacial edges with short peptide linkers. This setup limits structural flexibility, causing assembly failure; (*b*) $SBP1_{9.b}$ and $SBP2_{9.b}$ feature a flexible open-ended design: N/C-termini moved to interface vertices, placing interfacial edges at path ends. Central triangular edges are traversed once (low constraint), while pyramidal edges are traversed twice (high stability). This improves interfacial flexibility, allowing accurate cage assembly [10].

Further innovation emerged through covalent cyclization strategies [12] (See Figures 2 and 3). Subunit preorganization via split intein splicing enabled precise folding of tetrahedral designs, reducing maximum particle diameter $D_{\max}$ from 17.1 nm to 10.5 nm. This culminated in the trimeric SB24 octahedron, a 109-kDa architecture comprising 24 segments, where cyclization of the SB6 linker subunit induced significant conformational compaction ($D_{\max}$ from 25 nm to 9.4 nm), yielding the largest coiled-coil protein origami (CCPO) assembly to date.

These nanostructures such as bipyramidal nanocages, tetrahedral coils, and the SB24 octahedral complex exhibit a universal topological signature: interfacial edges (e.g., attachment faces) are traversed exactly once, while non-interfacial edges are traversed twice across their polyhedral frameworks. This pattern manifests in bipyramidal systems through dual closed walks (Figure 1), tetrahedral assemblies via paired trajectories (Figure 2), and culminates in the SB24 complex's three-walk configuration on an irregular octahedron (Figure 3).

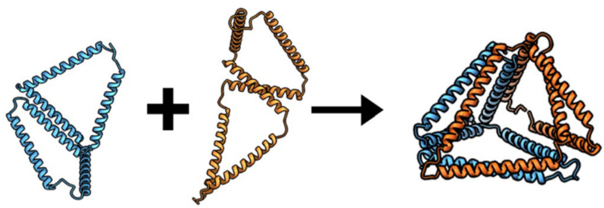


Figure 2. The bitrigon module structure of a two-chain coiled-coil protein is formed by two single-chain coiled-coil dimers. [12].

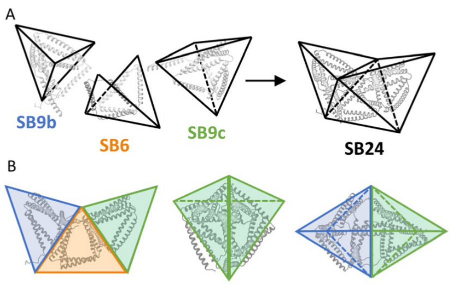


Figure 3. (A) The SB24 protein complex forms when two nine-segment subunits (SB9b and SB9c) and a six-segment peptide (SB6) interlock, depicted by geometric shapes; (B) The viewing angles of SB24 from left to right are side view, front view and top view.

In this paper, we employ a mathematical model to examine and define intricate cage-like formations formed by attaching faces of given polyhedra with the same degree, resulting in irregular polypeptide nanostructures. We investigate the oligomeric state of these self-assembled structures, specifically the number of their components, and analyze whether the constructed structures are stable or strongly stable. To support this, we extend the concept of double trace by introducing the semi-double trace, defined as a closed walk that traverses edges of Eulerian subgraphs exactly once and other edges twice, to explore the relationship between mathematics and polypeptide self-assembly. We further define the semi-strong trace to represent more stable configurations and provide conditions under which a semi-double trace is semi-strong. Additionally, we introduce double covers with multiple components that are strong if each component is semi-strong. Our model aims to elucidate the fundamental principles underlying these nanoscale structures and provide a theoretical foundation for the future design and synthesis of more complex biomimetic assemblies.

2 Preliminaries

2.1 Basic knowledge in graph theory

We will outline some basic terminology and results in graph theory that will be adopted for this paper. The graphs discussed in this paper are finite, connected, and simple except for a special case mentioned later. Some basic knowledge in graph theory can be found in [15]. Let G be a graph with vertex set $V(G)$ and edge set $E(G)$ respectively. Then the numbers of vertices and edges in G are denoted by $|V(G)|$ and $|E(G)|$, respectively. Suppose uv is an edge of G. Let $G - uv$ denote the graph obtained from G by deleting the edge uv.

Denote a **walk** in G by $W = v_0e_0v_1e_1 \cdots v_me_mv_{m+1}$, whose terms are alternatively vertices and edges and for every $i = 0, 1, \cdots, m$, where e_i denotes an edge connecting the vertices v_i and v_{i+1}. If $v_0 = v_{m+1}$, then the walk is called **closed**. The walk W is called a **trace** (or **path**) if it traverses each edge (or vertex) once, and the trace (or path) is called closed if W is closed. If W is a path, it can also be denoted by $v_0v_1 \cdots v_{m+1}$. A closed trace of G is called an **Euler tour** if it traverses each edge of G only once. If G contains an Euler tour, then we call G an **Euler graph** or **Eulerian**. It is known that G is Eulerian if and only if it contains no odd-degree vertices.

A graph G is cellularly embedded in a closed surface Σ if G is embedded in Σ and any connected component of $\Sigma - G$ is a 2-cell, referred to as a face. Such an embedding is also known as a **2-cell embedding** and can be realized through a combinatorial embedding with a rotation system. The **maximum genus** of G is known as the largest genus $g(\Sigma)$ for orientable surfaces Σ where G has a 2-cell embedding, denoted by $g_M(G)$. Then Euler's formula yields that $g_M(G) \leq \lfloor \frac{\beta(G)}{2} \rfloor$, where $\beta(G)$ $(= |E(G)| - |V(G)| + 1)$ represents the Betti number of G. Further, we call G **upper-embeddable** if the equality holds. The upper-embeddability implies the graph achieves maximum surface embedding genus, corresponding to minimal number of faces. Then we have the following theorem.

Theorem 1. *[16] Let G be a graph with even (or odd) Betti number. Then G is upper-embeddable if and only if it contains a spanning tree such that $G - E(T)$ contains no (or only one) odd component.*

If G is upper-embeddable and $\beta(G)$ is even, then this embedding is also called a **1-face embedding** and G is called **strictly upper-embeddable**.

2.2 Double traces and stable traces

Double trace has been a mathematical model for designing polypeptide structures [3, 10]. A closed walk of a graph G that traverses every edge of G twice is called a **double trace**. Let $W = w_0e_1w_1 \cdots e_lw_0$ be a double trace of G, with a length l. For any vertex v of G, let $M \subseteq N(v)$, where $N(v)$ represents the set of vertices adjacent to v. We say that W has an M**-repetition** at v if for all integers i such that $v_i = v$ and the pair $\{v_{i-1}, v_{i+1}\}$ is either contained in M or disjoint from M. Intuitively, whenever we enter v from a vertex in M, we also exit to a vertex in M from v. If an M-repetition at the vertex v satisfies $|M| = d$, then we call this repetition a d**-repetition** (or a repetition of order d), as shown in Figure

4. The M-repetition is called **trivial** if it satisfies $M = N(v)$ or $M = \emptyset$. A double trace is called **antiparallel** (or **parallel**) if any edge is traversed in opposite (or the same) directions. A double trace that has no nontrivial repetitions with order $\leq d$ is referred to as a d**-stable trace**. A **strong trace** is a double trace that has no nontrivial repetitions. For example, let $P = v_0e_0v_1$ be a path and $W_0 = v_0e_0v_1e_0v_0$ be a walk of P. Then W_0 is an antiparallel double trace of P and has a $\{v_0\}$ (or $\{v_1\}$)-repetition at the vertex v_1 (or v_0) which is also a trivial repetition. Thus W_0 is an antiparallel strong trace of P.

We consider a double trace of a graph to be stable as long as each vertex has no non-trivial repetitions without considering the directions of edge traversal, although this stability can be reinforced by parallel traversals that enhance edge rigidity and antiparallel traversals that promote vertex flexibility, mitigating steric hindrance.

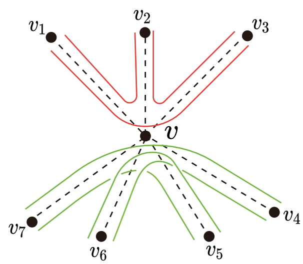


Figure 4. A $\{v_1, v_2, v_3\}$-repetition of order 3 (marked in red) and a $\{v_4, v_5, v_6, v_7\}$-repetition (marked in green) of order 4 at vertex v with degree 7.

By replacing each edge of G with a pair of antiparallel directed edges, we can apply Euler's theorem to derive the following result:

Proposition 2. *Any connected graph G admits an antiparallel double trace.*

The following theorem establishes a connection between the embedding of a graph and its strong trace.

Theorem 3. *[5] A graph G contains a strong antiparallel trace if and only if it admits a 1-face embedding in some orientable closed surface.*

3 Semi-double trace

In this section, we introduce semi-double trace and provide some basic properties. As a generalization of double trace, this concept reveals more connections between mathematics and the self-assembly of polypeptide nanostructures, as shown in Figure 1, 2, and 3.

Definition 1. For a graph G, a **semi-double trace** is a closed walk that traverses each edge of a nonempty subgraph $H \subseteq G$ exactly once and all other edges of G twice, where H is an Eulerian graph or a union of edge-disjoint cycles. When H consists of a single cycle, the semi-double trace is called **single**.

Note that: A crucial observation through the self-assembly of polypeptide structures indicates that the subgraph H must contain no odd degree vertices, ensuring the existence of a closed walk traversing each edge of H exactly once. Since if we replace any edge of G with a pair of undirected parallel edges, except for those edges that are traversed once, we can not obtain an Euler graph. If we suppose H is an empty graph, then the trace corresponds to the trivial case, making this definition a generalization of double traces.

Proposition 4. *Suppose G is a connected graph. Then G contains a semi-double trace if and only if it is not a tree.*

As illustrated in Figure 1, the component attached to each triangular bipyramid can be treated as a semi-double trace, where a triangular cycle is traversed once. As a more complex example, the semi-double traces (marked by solid lines) shown in Figure 5 traverse the edges marked in red (dashed lines) once. The subgraph induced by the red dashed lines can be viewed as consisting of edge-disjoint cycles.

Let W_H denote a semi-double trace of G where $H = \cup_{i=1}^{m} C_i$ and C_i $(i = 1, \cdots, m)$ are the edge-disjoint cycles of G such that W_H traverses the edges of each C_i once and other edges of G twice. Based on the stability and some significant effects on biology, we add a closed trace around each cycle C_i and let $W_{\widetilde{H}}$ denote this family of multi-component traces, such that any edge of G is traversed exactly twice. This construction maximizes

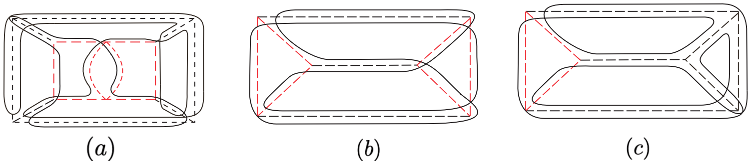


Figure 5. (*a*) A semi-double trace (marked by solid lines) traverses the edges on some Euler graph (marked by red dashed lines) once; (*b*) A semi-double trace (marked by solid lines) traverses the edges on two edge-disjoint cycles (marked by red dashed lines) once; (*c*) A single semi-double trace (marked by solid lines).

the number of trace components, as shown in Figure 6. Then any edge in G is traversed twice by $W_{\widetilde{H}}$, which is also called **the induced traces** of W_H. The number of components of $W_{\widetilde{H}}$ is $m+1$. An examples is shown in Figure 7.

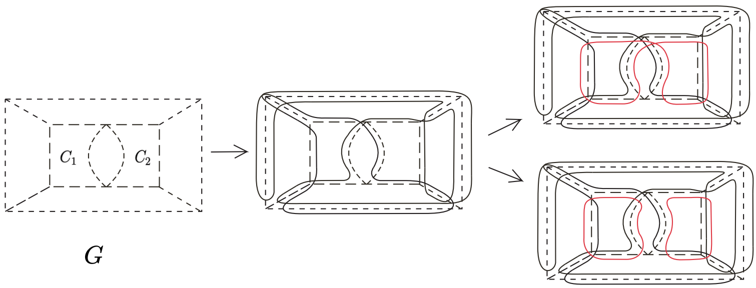


Figure 6. A semi-double trace of G is formed by traversing edges in the cycles C_1, C_2 once, and the other edges twice. We can add one or two closed traces around the edges that are traversed once and form the multi-component traces such that any edge of G is traversed twice.

As a special case of semi-double trace, let $W_\emptyset$ represent a double trace. Since the definition of repetition is determined by the local behavior of a double trace, we can also extend the definition of repetition to the induced traces $W_{\widetilde{H}}$ of a semi-double trace W_H, which also traverses every edge twice . We define the repetition of $W_{\widetilde{H}}$ in the same way as that of a double trace $W_\emptyset$ (See Section 2.2), so we omit the description of the definition.

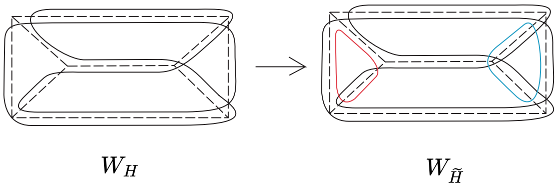


Figure 7. W_H is a semi-double trace; each side of the leftmost and rightmost triangles (edge-disjoint cycles) of G is covered only once. Then $W_{\widetilde{H}}$ can be obtained by adding red and blue traces.

Furthermore, we can also define a vertex figure of $W_{\widetilde{H}}$, denoted by $F_{v,W_{\widetilde{H}}}$ (See Definition 3). The definition of "semi-strong" is as follows.

Definition 2. A semi-double trace W_H of a graph G is called **semi-d-stable** if its induced traces $W_{\widetilde{H}}$ is d-stable. In particular, W_H is called **semi-strong** if $W_{\widetilde{H}}$ is strong.

Definition 3. Let W_H be a semi-double trace of a graph G and v be a vertex of G. Let $W_{\widetilde{H}}$ be the induced traces of W_H. We use $E(v)$ as the vertex set to construct the vertex figure of W_H (or $W_{\widetilde{H}}$) at v, where the edges e and e' in $E(v)$ are considered adjacent if they are consecutive along W_H (or $W_{\widetilde{H}}$). This vertex figure is denoted as F_{v,W_H} (or $F_{v,W_{\widetilde{H}}}$).

For a double trace $W_\emptyset$, any vertex figure $F_{v,W_\emptyset}$ of $W_\emptyset$ is 2-regular. Note that the vertex figure of the induced traces, $F_{v,W_{\widetilde{H}}}$, of a semi-double trace W_H is also 2-regular. For example, the vertex figure of the vertex v shown in Figure 8 is a cycle and is denoted by $e_1e_2\cdots e_ne_1$.

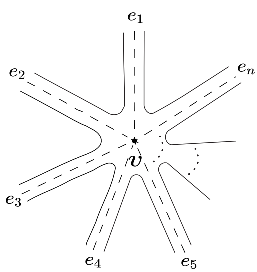


Figure 8. A trivial repetition at the vertex v where $e_1, \cdots, e_n$ are the edges incident with v.

Similarly as the version of double trace, we have the following Theorem 5. We omit the proof and recommend readers to refer to [5].

Theorem 5. *Let G be a graph and W_H be a semi-double trace of G. Let $W_{\widetilde{H}}$ be the induced traces of W_H. Then $W_{\widetilde{H}}$ is strong if and only if any vertex figure $F_{v,W_{\widetilde{H}}}$ is a cycle.*

In the following, we characterize semi-strong traces.

Lemma 1. *Let G be a graph, and let W_H denote a semi-double trace of G. Let C be a cycle of G where W_H traverses every edge once. Suppose $v \in V(C)$ and e_v, e'_v are the two edges incident with v in C. Then the vertex figure F_{v,W_H} is a subgraph of $F_{v,W_{\widetilde{H}}} - e_ve'_v$.*

Proof. According to the construction, $F_{v,W_{\widetilde{H}}}$ must contain an edge of $e_ve'_v$, and F_{v,W_H} is a subgraph of $F_{v,W_{\widetilde{H}}}$. If $e_ve'_v \notin E(F_{v,W_H})$, then the consequence holds. If $e_ve'_v \in E(F_{v,W_H})$, since W_H traverses the edges of C once, then F_{v,W_H} contains no parallel edges of $e_ve'_v$. Since $e_ve'_v \in E(F_{v,W_H})$ and F_{v,W_H} is a subgraph of $F_{v,W_{\widetilde{H}}}$, then according to the construction, we have $F_{v,W_{\widetilde{H}}}$ contains a 2-cycle with parallel edges of $e_ve'_v$. Thus $F_{v,W_{\widetilde{H}}} - e_ve'_v$ contains an edge of $e_ve'_v$ and then F_{v,W_H} is a subgraph of $F_{v,W_{\widetilde{H}}} - e_ve'_v$. ■

Theorem 6. *Let G be a graph and W_H be a semi-double trace of G, where $H = \cup_{i=1}^{m} C_i$ for any $E(C_i) \cap E(C_j) = \emptyset$ $(i \neq j)$. If W_H is semi-strong, then we have,*

(1) for $v \in V(H)$, every vertex figure F_{v,W_H} is a path or disjoint union of some paths;

(2) for $v \notin V(H)$, every vertex figure F_{v,W_H} is a single cycle.

Proof. The result of (2) in this theorem is obvious according to Theorem 5. Then we only need to prove (1).

Suppose W_H is semi-strong, then its induced traces $W_{\widetilde{H}}$ is strong. For $v \in V(H)$, then the vertex figure $F_{v,W_{\widetilde{H}}}$ must be a cycle, denoted by $e_1e_2\cdots e_ke_1$, where e_i $(i = 1, 2, \cdots, k)$ are the vertices of $F_{v,W_{\widetilde{H}}}$ and $k = |V(F_{v,W_{\widetilde{H}}})|$. Since $v \in V(H)$, then there exsits a C_i for some integer $1 \leq i \leq m$ such that $v \in V(C_i)$. Let e_v, e'_v be the two edges incident with v in C_i. Then $e_ve'_v$ is an edge of $F_{v,W_{\widetilde{H}}}$. Without loss of generality, suppose $e_v e'_v = e_1e_2$. Then according to Lemma 1, F_{v,W_H} is a subgraph of $F_{v,W_{\widetilde{H}}} - e_1e_2$ which is the path $e_2 \cdots e_k e_1$. Thus F_{v,W_H} is a path or disjoint union of some paths. ■

An example is shown in Figure 9. The left figure is the semi-strong trace W_H at v whose vertex figure (disjoint union of two paths) is represented in the right figure.

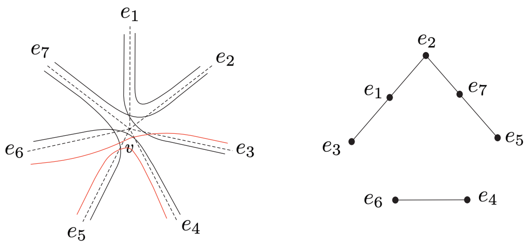


Figure 9. The left figure is the semi-strong trace W_H (marked by black solid lines) at v and the part of closed traces (marked by red solid lines) around cycles in H. Its vertex figure F_{v,W_H} (disjoint union of two paths) is represented in the right figure.

Note that: The converse stated in Theorem 6 is not true, but holds when W_H is a single semi-double trace. A counterexample is shown in Figure 10 where every vertex figure as indicated in condition (1) is a disjoint union of two paths, but the repetition at v of its induced traces is not trivial.

For a single semi-double trace, we have the following theorem.

Theorem 7. *Let G be a graph and W_H be a single semi-double trace of G, where H is a cycle of G. Then W_H is semi-strong if and only if any vertex figure of W_H at a vertex of H is a path, and other vertex figures are single cycles.*

Proof. According to Theorem 5 and 6, we only consider the vertices of the cycle H. Let v be a vertex of H and e_v, e'_v be the two edges incident with v in C. If W_H is semi-strong, then according to Theorem 6, F_{v,W_H} is a path or disjoint union of some paths. The induced traces $W_{\widetilde{H}}$ is strong, and the vertex figure $F_{v,W_{\widetilde{H}}}$ is a cycle which contains the edge $e_v e'_v$. Let

$E(v)$ be the edge set which contains all edges incident with v in G. Then any edge in $E(v) - \{e_v, e'_v\}$ is traversed twice when W_H traverses. Thus $F_{v,W_H} = F_{v,W_{\widetilde{H}}} - e_v e'_v$, which is a path with e_v, e'_v as its endpoints.

Suppose for any $v \in V(H)$, F_{v,W_H} is a path, other vertex figures of W_H are single cycles. According to the construction, $F_{v,W_{\widetilde{H}}} = F_{v,W_H} + e_v e'_v$. Further, $F_{v,W_{\widetilde{H}}}$ must be 2-regular, then the endpoints of F_{v,W_H} must be e_v and e'_v. Thus $F_{v,W_{\widetilde{H}}}$ is a cycle. This implies every vertex figure of $W_{\widetilde{H}}$ is a cycle and then W_H is semi-strong. ■

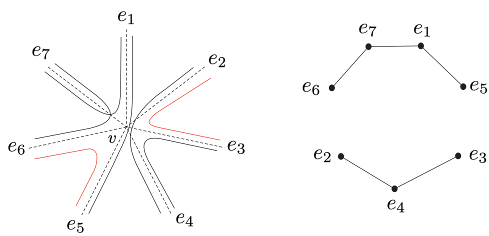


Figure 10. The right figure shows the vertex figure of a semi-double trace (marked by black solid lines) at v as indicated in the left figure. However, by considering its induced traces, the repetition at v is not trivial because if we cap off the cycles along $e_5, v, e_6, \cdots$ and $e_2, v, e_3, \cdots$, the vertex figure on v contains two cycles: $e_6e_7e_1e_5e_6$ and $e_2e_3e_4e_2$.

4 A strong stable model of self-assembly polypeptide structures

The intricate cage-like nanostructures observed in recent experiments—such as the two-chain bipyramidal nanocage (Figure 1), the two-chain tetrahedral coiled-coil (Figure 2), and the 24-segment SB24 octahedral complex (Figure 3)—demonstrate that polypeptide self-assembly inherently relies on multi-component closed walks traversing polyhedral edges with specific frequencies (once for interfacial edges, twice otherwise). To mathematically characterize such stability, we extend the concept of strong traces to multi-component systems. Here, we define a generalized version of double trace called double cover (Definition 4) and prove that its strong stability (Theorem 8) directly mirrors the biological stability observed in these structures.

The number of trace components in a double cover can be more than 3. Crucially, by constructing a strong double cover from semi-strong traces (Theorems 9-10), we establish a rigorous framework to design polypeptide cages with programmable oligomeric states and guaranteed topological stability.

Definition 4. Let $\overline{W}$ be a family of closed traces of a graph G such that $\overline{W}$ traverses every edge of G twice. Then $\overline{W}$ is called a **double cover** of G.

Note that: Any graph (including a disconnected graph) contains a double cover according to Proposition 2. A double cover with one component is also referred to as a double trace, thus also generalizing the concept of double trace. The induced traces of a semi-double trace, as mentioned in the previous section, is also a double cover.

For an edge e of G, if $\overline{W}$ traverses e in opposite (or the same) directions, then e is called antiparallel (or parallel). A double cover $\overline{W}$ is called **antiparallel** (or **parallel**) if any edge of G is traversed in an antiparallel (or parallel) manner.

We can also extend the definition of repetition to double cover. Then a double cover is called **strong** if it has no nontrivial repetitions. In [5], the authors established a relation between the double trace W of a connected graph G and a vertex figure at any vertex. Specifically, the necessary and sufficient condition of W to be strong is that any vertex figure $F_{v,W}$ forms a cycle. Actually, the number of components in the definition of $F_{v,W}$ does not need to be restricted to one, it can have multi components. As a direct consequence, we have Definition 5 and Theorem 8.

Definition 5. Let $\overline{W}$ be a double cover of a graph G and v be a vertex of G. We use $E(v)$ as the vertex set to construct the vertex figure at v of $\overline{W}$, where the edges e and e' in $E(v)$ are considered adjacent if they are consecutive along $\overline{W}$. This vertex figure is denoted as $F_{v,\overline{W}}$.

Theorem 8. *Let $\overline{W}$ be a double cover of a graph G and v be a vertex of G. Then $\overline{W}$ is strong if and only if any vertex figure $F_{v,\overline{W}}$ forms a cycle.*

Suppose W_{C_1} and W_{C_2} are two single semi-double trace of graphs G_1 and G_2 respectively. If $|E(C_1)| = |E(C_2)| = k$, let $C_i = v_1^i v_2^i \cdots v_k^i v_1^i$ where

v_j^i $(j = 1, 2, \cdots, k)$ are the vertices around C_i $(i = 1, 2)$. A new graph can be obtained by identifying C_1 and C_2 such that $v_j^1 = v_j^2$ $(j = 1, \cdots, k)$, forming a graph called **cover graph** of the ordered pair (G_1, G_2). Note that there are different ways to attach the graphs based on the ordering of the vertices in the presentation of C_i. In this new cover graph, W_{C_1} and W_{C_2} form a double cover. An example is shown in Figure 11 where the graph on the right is obtained by identifying the two cycles marked in yellow. Consequently, W_{C_1} and W_{C_2} form a double cover on the resulting graph, marked in red and blue. Then we have the following theorem.

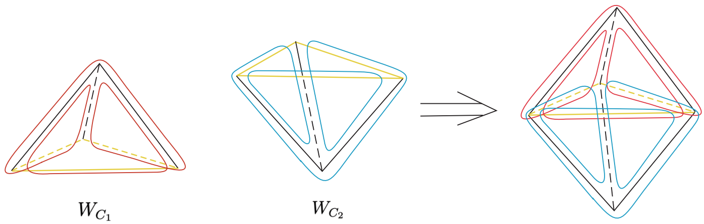


Figure 11. W_{C_1} and W_{C_2} are single semi-double traces (marked in red and blue). They traverse the edges of C_1 and C_2 (marked in yellow) respectively once, and the other edges twice. These traces are fabricated to be a double cover with two components; The cover graph is represented by the black and yellow graph in the rightmost.

Theorem 9. *Suppose W_{C_1} and W_{C_2} are two single semi-strong traces of graphs G_1 and G_2 respectively, with $|E(C_1)| = |E(C_2)| = k$. Then the double cover $\overline{W}$ on a cover graph of (G_1, G_2) is strong, analogous to viral capsid stability from symmetric protein interactions.*

Proof. Let $C_i = v_1^i v_2^i \cdots v_k^i v_1^i$ $(i = 1, 2)$ where v_j^i $(j = 1, 2, \cdots, k)$ are the vertices around C_i. It is sufficient to prove that any vertex figure of v_j^i for the double cover $\overline{W}$ is a cycle. Since W_{C_i} $(i = 1, 2)$ are semi-strong, then according to Theorem 7, the vertex figure of v_j^i in W_{C_i} is a path, denoted by $P_i = e_1^i e_2^i \cdots e_{m_i}^i$, where $m_i = d_{G_i}(v_j^i) - 1$. The endvertices of P_i, e_1^i and $e_{m_i}^i$, are the two edges incident with v_j^i in C_i. Then according to the construction of the cover graph, we have

$$e_1^1 = e_1^2, \quad e_{m_1}^1 = e_{m_2}^2,$$

then the vertex figure $F_{v_j^i,\overline{W}}$ can be denoted by

$$P_1P_2 = e_1^1e_2^1\cdots e_{m_1}^1(=e_{m_2}^2)e_{m_2-1}^2\cdots e_1^2.$$

Since any edge $e_j^1 \in E(G_1)$ and $e_j^2 \in E(G_2)$, and $e_1^1 = e_1^2$, then P_1P_2 is a cycle. Thus $\overline{W}$ is a strong double cover. ■

We then generalize the above Theorem 9. Let W_{C_i} be a single semi-double trace of G_i $(i = 1, 2, \cdots)$ where C_i is a cycle in G_i. Suppose $|E(C_i)| = k_i$, and let $C_i = v_1^iv_2^i\cdots v_{k_i}^iv_1^i$ where v_j^i $(j = 1, \cdots, k_i)$ are the vertices around C_i.

In the first step, we need to identify two cycles or paths of G_1 and G_2. If $k_1 = k_2$ and we identify C_1 and C_2, we obtain a newly constructed graph as described in the previous construction of the cover graph of the ordered pair (G_1, G_2). We denote this new graph by $G^{(1)}$, which has a double cover with 2-components, and the procedure terminates. Alternatively, we may continue the procedure and identify two paths P_1 and P_2 with the same length $l \leq min\{k_1, k_2\}$ in C_1 and C_2 respectively, and denote the new constructed graph by $G^{(1)}$ as well. Suppose $P_i = v_1^iv_2^i\cdots v_l^i$ $(i = 1, 2)$ and let $v_j^1 = v_j^2$ $(j = 1, 2, \cdots, l)$. Then we obtain a cycle

$$C' = v_l^1v_{l+1}^1\cdots v_{k_1}^1v_1^1(=v_1^2)v_{k_2}^2v_{k_2-1}^2\cdots v_{l+1}^2v_l^2(=v_l^1),$$

and the edges of C' are traversed once by both W_{C_1} and W_{C_2} on $G^{(1)}$. This implies the edges of $G^{(1)}$ which are traversed once form a cycle. Then we continue the procedure.

In the second step, we construct $G^{(2)}$ in the same manner by identifying two cycles or paths with the same length in C' of $G^{(1)}$ and C_3 of G_3 respectively. If $|E(C')| = |E(C_3)|$ and the two cycles are identified, each edge of $G^{(2)}$ is traversed twice, and the procedure terminates. Otherwise, if we identify two paths as described earlier, some edges forming a cycle in $G^{(2)}$ will be traversed only once, and the procedure continues.

We repeat this procedure until we reach the step $n-1$ for some integer n. At this step, using the same method, we identify the cycle in $G^{(n-2)}$ which are traversed once and the cycle C_n in G_n. Note that these two cycles must have the same length. This results in the construction of $G^{(n-1)}$ which contains a double cover composed of $(W_{C_1}, W_{C_2}, \cdots, W_{C_n})$. The graph $G^{(n-1)}$ is also called a cover graph of $(G_1, G_2, \cdots, G_n)$.

An example of a cover graph is shown in Figure 12 where the cover graph is obtained by identifying C_1, C_2, and C_3 (marked in yellow) to form a new graph $G^{(2)}$. Simultaneously, the single semi-double traces W_{C_1}, W_{C_2}, and W_{C_3} form a double cover on $G^{(2)}$, marked in red, green, and purple respectively. Then we have the following theorem.

Theorem 10. *Suppose W_{C_i} is a single semi-strong trace of a graph G_i $(i = 1, 2, \cdots, m)$ where C_i is a cycle in G_i, and $G^{(m-1)}$ is a cover graph of $(G_1, \cdots, G_m)$. Let $\overline{W}$ be the corresponding double cover of $G^{(m-1)}$. Then $\overline{W}$ is strong.*

Proof. For any vertex $v \in V(G^{(m-1)})$, If $v \notin V(C_i)$, $i = 1, 2, \cdots, m$, then the vertex figure $F_{v,\overline{W}} = F_{v,W_{C_i}}$ is a cycle. If $v \in V(C_i)$ for some i, let $\sum_{k=1}^{m} \mathbb{1}_{v \in V(C_k)}$ denote the total number of occurrences of the element v across all $V(C_1)$ to $V(C_m)$, where $\mathbb{1}_{v \in V(C_k)}$ is an indicator function that is 1 if the element $v \in V(C_k)$, and 0 otherwise, i.e., the multiplicity of v in the cover graph assembly. Then $2 \leq \sum_{k=1}^{m} \mathbb{1}_{v \in V(C_k)} \leq m$. We assume $\sum_{k=1}^{m} \mathbb{1}_{v \in V(C_k)} = q$, then there exist q edge subsets incident with v, which are traversed by q different single semi-strong traces respectively. Without loss of generality, the traces are denoted by $W_{C_1}, \cdots, W_{C_q}$. A similar discussion to that in Theorem 9 shows that the vertex figure $F_{v,\overline{W}}$ is composed of the vertex figures $F_{v,W_{C_1}}, F_{v,W_{C_2}}, \cdots, F_{v,W_{C_q}}$, which are paths that form a cycle in an end-to-end manner. Consequently, $\overline{W}$ is a strong double cover. ■

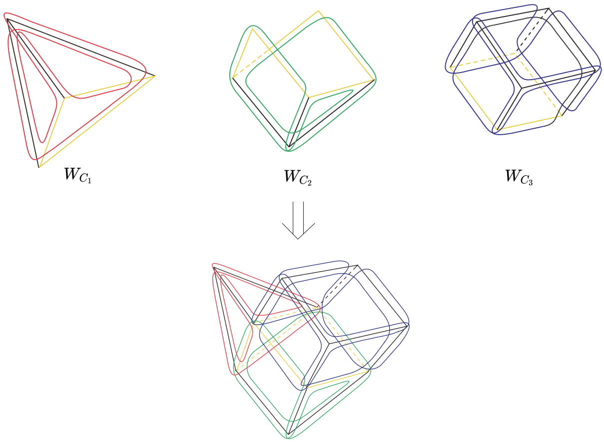


Figure 12. W_{C_1}, W_{C_2} and W_{C_3} are single semi-double traces (marked in red, green and purple). They pass through cycles C_1, C_2 and C_3 (marked in yellow) respectively once, and the other edges twice. These traces are created as a double cover with three components.

Stability in polypeptide structures refers to the robust interlocking configuration where each edge is traversed twice by a single chain, forming coiled-coil dimers, and each vertex has no non-trivial repetitions. These coiled-coil dimers are formed and interlocked to create a stable structure. The double cover model directly reflects the stoichiometry of polypeptide chains, such as the case like the SB24 complex (matching $n = 3$ components, which requires three polypeptide chains in self-assembled cages, with each component in the double cover corresponding to one chain). Building on this, Theorem 10 implies the double cover as shown in Figure 12 is a strong double cover, and based on that, we can construct a strong stable model for the structure as shown in Figure 3.

5 Conclusion

This paper presents a graph-theoretic framework utilizing semi-double traces and multi-component double covers to model the self-assembly of coiled-coil protein origami (CCPO) cages, inspired by recent experimental advancements such as the two-chain triangular bipyramids, tetrahedral coils, and the three-chain SB24 octahedral complex. CCPO cages provide robust interlocking stability and enable potential applications such as drug delivery and biosensing. These cage-like structures, formed by single-chain coiled-coils on polyhedral graphs, feature interfacial edges that are traversed once to support conformational adaptability and non-interfacial edges that are traversed twice for structural reinforcement, serving as test cases for our model.

The mathematical model focuses on utilizing semi-strong traces and double covers. For graphs $G_1, G_2, \ldots, G_n$, we construct the cover graph G^{n-1}, representing an n-chain cage where each chain on G_i follows a single semi-double trace W_{C_i}. When this structure, typically a polyhedron, is designed as a coiled-coil protein origami, the number of chains is n. According to our model, the origami corresponds to a double cover with n components, where each single-chain coiled-coil on G_i corresponds to a single semi-double trace. If every semi-double trace is semi-strong, the resulting double cover with n components is strong. Strong double covers ensure the stability of these protein cages by preventing nontrivial repetitions that could lead to assembly failures.

A key innovation is the integration of experimental insights from [10] and [12] regarding N- and C-termini positioning into our model. Lapenta et al. demonstrated that the successful assembly of the $\mathrm{SBP1}_{9.\mathrm{b}}/\mathrm{SBP2}_{9.\mathrm{b}}$ bipyramid requires N- and C-termini at interaction vertices, in contrast to the failed $\mathrm{SBP1}_{9.\mathrm{a}}/\mathrm{SBP2}_{9.\mathrm{a}}$ design where non-interface termini led to rigidity. In our model, the vertices of C_i are interpreted as the positions of the N- and C-termini of the respective polypeptide chains. A design with termini at interface vertices yields a single-cycle vertex figure, ensuring a semi-strong trace and enhanced stability. This is validated by SB24 complex, where cyclization pre-organizes chains, aligning with our semi-strong trace condition for the 109-kDa structure.

Compared to single-component models, our approach determines oligomeric states from component counts, extends to branched nanostructures via semi-double traces, and provides a stability framework via cyclic vertex figures. This prescriptive tool guides the design of programmable biomimetic nanomaterials, such as the SB24 octahedron, with predictable stability.

Future work will quantify cyclization energy and linker effects to further connect graph theory with synthetic biology.

***Acknowledgment*:** S. Liu is supported by National Natural Science Foundation of China (No. 12201471). J. Ning is supported by National Natural Science Foundation of China (No. 12171402). X.-S. Cheng is supported by Huizhou City Basic Education educational science research project (No. 2023hzkt193), Teaching Quality and Teaching Reform Project of Huizhou University (No. 158 [2023]).

References

[1] D. Bhatia, S. Mehtab, R. Krishnan, S. S. Indi, A. Basu, Y. Krishnan, Icosahedral DNA nanocapsules by modular assembly, *Angew. Chem. Int. Ed.* **48** (2009) 4134–4137.

[2] J. H. Chen, N. C. Seeman, Synthesis from DNA of a molecule with the connectivity of a cube, *Nature* **350** (1991) 631–633.

[3] X.-S. Cheng, Q. Deng, Y. Diao, Constructions of DNA and polypeptide cages based on plane graphs and odd crossing π-junctions, *Appl. Math. Comput.* **443** (2023) 127773–127788.

[4] X.-S. Cheng, Y. Diao, Configurations of DNA cages based on plane graphs and vertex junctions, *J. Phys. A: Math. Theor.* **53** (2020) 395601–395622.

[5] G. Fijavž, T. Pisanski, J. Rus, Strong traces model of self-assembly polypeptide structures, *MATCH Commun. Math. Comput. Chem.* **71** (2014) 199–212.

[6] R. P. Goodman, R. M. Berry, A. J. Turberfield, The single-step synthesis of a DNA tetrahedron, *Chem. Commun.* **12** (2004) 1372–1373.

[7] H. Gradišar, S. Božič, T. Doles, D. Vengust, I. Hafner-Bratkovič, A. Mertelj, B. Webb, A. Šali, S. Klavžar, R. Jerala, Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments, *Nat. Chem. Biol.* **9** (2013) 362–366.

[8] Y. He, T. Ye, M. Su, C. Zhang, A. E. Ribbe, W. Jiang, C. Mao, Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra, *Nature* **452** (2008) 198–201.

[9] S. Klavžar, J. Rus, Stable traces as a model for self-assembly of polypeptide nanoscale polyhedrons, *MATCH Commun. Math. Comput. Chem.* **70** (2013) 317–330.

[10] F. Lapenta, J. Aupič, M. Vezzoli, Ž. Strmšek, S. Da Vela, D. I. Svergun, J. M. Carazo, R. Melero, R. Jerala, Self-assembly and regulation of protein cages from pre-organized coiled-coil modules, *Nat. Commun.* **12** (2021) 939–951.

[11] W. M. Shih, J. D. Quispe, G. F. Joyce, A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron, *Nature* **427** (2004) 618–621.

[12] J. Snoj, F. Lapenta, R. Jerala, Preorganized cyclic modules facilitate the self-assembly of protein nanostructures, *Chem. Sci.* **15** (2024) 3673–3687.

[13] J. Wang, X. Jin, M. Ji, On the existence of F-strong trace of a graph when F induces a forest, *MATCH Commun. Math. Comput. Chem.* **77** (2017) 799–812.

[14] J. Wang, G. Hu, M. Ji, Almost parallel strong trace model of self-assembly polypeptide nanostructure, *MATCH Commun. Math. Comput. Chem.* **77** (2017) 783–798.

[15] D. B. West, *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, 1996.

[16] N. H. Xuong, How to determine the maximum genus of a graph, *J. Comb. Theory Ser. B* **26** (1979) 217–225.

[17] C. Zhang, S. H. Ko, M. Su, Y. Leng, A. E. Ribbe, W. Jiang, C. Mao, Symmetry controls the face geometry of DNA polyhedra, *J. Am. Chem. Soc.* **131** (2009) 1413–1415.

[18] J. Zimmermann, M. P. J. Cebulla, S. Mönninghoff, G. von Kiedrowski, Self-assembly of a DNA dodecahedron from 20 trisoligonucleotides with C_{3h} linkers, *Angew. Chem. Int. Ed.* **47** (2008) 3626–3630.