

# Discrete Exterior Calculus of Proteins and Their Cohomology

Naoto Morikawa 

Genocript, Zama, Japan

Email: nmorika@genocript.com

**How to cite this paper:** Morikawa, N. (2022) Discrete Exterior Calculus of Proteins and Their Cohomology. *Open Journal of Discrete Mathematics*, 12, 47-63. <https://doi.org/10.4236/ojdm.2022.123004>

**Received:** May 13, 2022

**Accepted:** July 10, 2022

**Published:** July 13, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

This paper proposes a novel application of cohomology to protein structure analysis. Since proteins interact each other by forming transient protein complexes, their shape (e.g., shape complementarity) plays an important role in their functions. In our mathematical toy models, proteins are represented as a loop of triangles (2D model) or tetrahedra (3D model), where their interactions are defined as fusion of loops. The purpose of this paper is to describe the conditions for loop fusion using the language of cohomology. In particular, this paper uses cohomology to describe the conditions for “allosteric regulation”, which has been attracted attention in safer drug discovery. I hope that this paper will provide a new perspective on the mechanism of allosteric regulation. Advantages of the model include its topological nature. That is, we can deform the shape of loops by deforming the shape of triangles (or tetrahedra) as long as their folded structures are preserved. Another advantage is the simplicity of the “allosteric regulation” mechanism of the model. Furthermore, the effect of the “post-translational modification” can be understood as a resolution of singularities of a flow of triangles (or tetrahedra). No prior knowledge of either protein science, exterior calculus, or cohomology theory is required. The author hopes that this paper will facilitate the interaction between mathematics and protein science.

## Keywords

Discrete Differential Geometry, Protein Structure Analysis, Cohomology Class, Exterior Derivative, Allosteric Regulation

---

## 1. Introduction

Proteins are hardworking macromolecules that perform a variety of functions in

cells. Proteins are made up of chains of 50 to 2000 amino acids, usually folded into a well-defined three-dimensional structure. Since proteins interact each other by forming transient protein-protein complexes, their shape (such as shape complementarity at the interface) plays an important role in their functions.

In this paper, three mathematical models of proteins are presented, *i.e.*, the loop model, the flow model, and the cohomological model in this order (**Figure 1**).

First, in Section 2, we explain the problem intuitively using a loop model. In particular, we give the intuitive definition of protein interactions and allosteric triplets. For example, “protein interaction is defined as fusions of loops”.

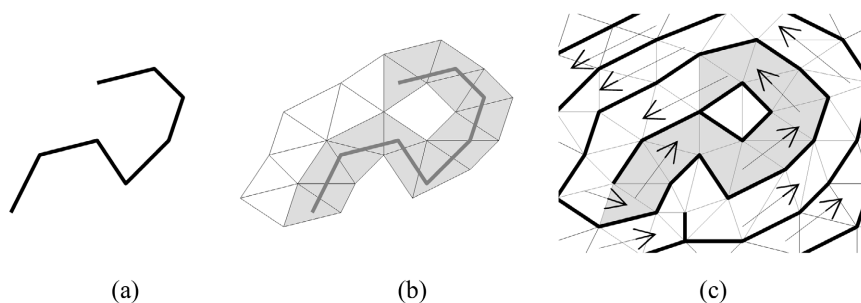
Next, in Section 3, we consider conditions for protein interaction and allosteric regulation using the flow model. The flow model is a differential geometric model [1], where proteins are represented as a closed trajectory of a flow of triangles (2D model) or tetrahedra (3D model). Then, we can define conditions for protein interaction and allosteric regulation using the language of differential geometry [2]. For example, “two proteins interact if the corresponding local flow is integrable”.

Finally, in Section 4, we rephrase the conditions obtained in Section 3 using the cohomological model. For example, “two proteins interact if the cohomology class of the corresponding vector field is zero”. Cohomology classes of vector fields are defined using exterior derivative operators.

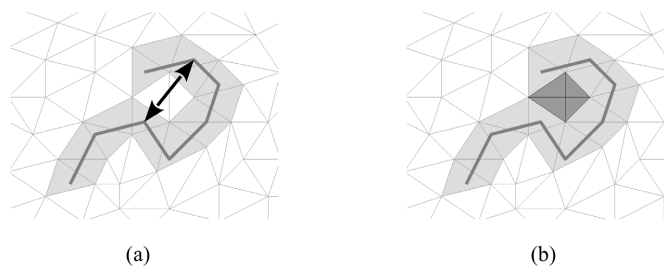
“Allosteric regulation” and “post-translational modification” are briefly introduced in Subsections 2.3 and 3.6, respectively. In the flow/cohomological model, “allosteric regulation” corresponds to the integrability of a flow, and “post-translational modification” corresponds to a resolution of singularities of a flow (**Figure 2**).

As for previous works, the author is unaware of any other geometrical models of allosteric regulation nor any applications of cohomology to protein structure analysis. For an overview of history of allostery, see [3]. For an application of cohomology to biological time series, see [4].

For the sake of clarity, we mainly consider loops of triangles and triangular meshes. In the following, we denote the set of all integers by  $\mathbb{Z}$ .



**Figure 1.** The mathematical models. (a) Protein (Polyline); (b) The loop model; (c) The flow/cohomological model.



**Figure 2.** Advantages of the flow/cohomological model. (a) Integrability of a flow. The arrow indicates a conflict between two distal triangles across two “not-form-a-loop” triangles; (b) Resolution of singularities of a flow. By splitting each of the “not-form-a-loop” triangles into two, a loop of length four (dark grey) is obtained.

## 2. Loop Model of Proteins

In the loop model (**Figure 1(b)**), amino acid sequences are represented as a closed chain of triangles (2D model) or tetrahedra (3D model).

### 2.1. Loops and Their Normal Edges

A *chain  $u$  of triangles* is a series of triangles connected by a common edge, *i.e.*, for some interval  $I \subset \mathbb{Z}$ ,

$$u = \{t[i] (i \in I) \mid t[i] \neq t[j] \text{ if } i \neq j \\ t[i] \text{ and } t[i+1] \text{ intersect at a common edge}\}. \tag{1}$$

A closed chain  $u$  is called a *loop*.

Let  $u = \{t[i]\}$  be a chain of triangles. The *underlying mesh*  $M(u)$  of  $u$  is a triplet defined by

$$M(u) := \langle T(u), E(T(u)), V(T(u)) \rangle, \tag{2}$$

where

$$\begin{cases} T(u) := \{t \mid t \in u\}, \\ E(T(u)) := \text{the set of all edges of the triangles in } T(u), \\ V(T(u)) := \text{the set of all vertices of the triangles in } T(u). \end{cases} \tag{3}$$

In this paper, we consider only the chains such that (1) two triangles in  $T(u)$  do not intersect except at an edge, and (2) the edges in  $E(T(u))$  are shared by at most two triangles in  $T(u)$ . In particular,

$$E(T(u)) = E_1(T(u)) \cup E_2(T(u)) \text{ (disjoint union)}, \tag{4}$$

where  $E_k(X)$  denotes the set of edges in  $E(X)$  shared by  $k$  triangles in  $X$ . Edges in  $E_1(T(u))$  are called the *boundary edges* of  $M(u)$ .

Let  $u = \{t[i]\}$  be a chain of triangles. Let  $t \in u$ . The *normal edge*  $N(t)$  of  $t$  is the edge that is not shared with the adjacent triangles along  $u$ . (In **Figure 1(c)**, the normal edges are drawn with thick line segments.) If  $u$  is a loop, each triangle in  $u$  has exactly one normal edge. If  $u$  is not a loop, the endpoint triangles have two normal edges. If  $u$  consists of one triangle, say  $u = \{t\}$ ,  $t$  has three

normal edges. We denote the set of all normal edges of triangles in  $u$  by  $N(u)$ :

$$N(u) := \{N(t) \mid t \in u\}. \tag{5}$$

**Remark 2.1.** The normal edge to a chain is a discrete version of the “normal vector” to a curve.

The pair  $\langle M(u), N(u) \rangle$  of  $M(u)$  and  $N(u)$  is called the *local flow* of  $u$ . Note that we can recover  $u$  uniquely from  $\langle M(u), N(u) \rangle$  by connecting triangles in  $M(u)$  along the edges in  $N(u)$ , *i.e.*,

**Lemma 2.2.** Let  $u = \{t[i]\}$  be a chain of triangles. Then, there is a one-to-one correspondence

$$u \leftrightarrow \langle M(u), N(u) \rangle. \tag{6}$$

In the same way, a *chain  $u$  of tetrahedra* is defined as a series of tetrahedra connected by a common face. The *normal edge* of a tetrahedra in a trajectory is defined as the edge that is not shared with two adjacent tetrahedra along the trajectory.

### 2.2. Loop Interaction

Let  $u_a$  and  $u_b$  be two chains of triangles. Binary set operations between  $M(u_a)$  and  $M(u_b)$  are defined by

$$\begin{aligned} &M(u_a) \square M(u_b) \\ &:= \langle T(u_a) \square T(u_b), E(T(u_a) \square T(u_b)), V(T(u_a) \square T(u_b)) \rangle, \end{aligned} \tag{7}$$

where  $\square$  stands for a set operation such as  $\cup, \cap, \setminus$ , and others. Binary relations between  $M(u_a)$  and  $M(u_b)$  are defined by

$$M(u_a) \square M(u_b) \text{ if and only if } T(u_a) \square T(u_b) \tag{8}$$

where  $\square$  stands for a binary relation such as  $\subset, \supset$ , and others.

**Remark 2.3.** In this paper, we suppose that  $E_k(T(u_a) \square T(u_b) \square \dots) = \emptyset$  for  $k > 2$ . For example, we consider  $M(u_a) \cup M(u_b)$  only for  $u_a$  and  $u_b$  with no overlap.

**Definition 2.4. (Reaction Intermediate)** Let  $u_a, u_b, \dots, u_c$  be loops of triangles.  $M(u_a) \cup M(u_b) \cup \dots \cup M(u_c)$  is called *loop-transitive* if there are loops  $u, u_1, u_2, \dots, u_m$  such that

$$\begin{cases} T(u) \subset T(u_a) \cup T(u_b) \cup \dots \cup T(u_c), \\ E(T(u)) \supset E_1(T(u_a) \cup T(u_b) \cup \dots \cup T(u_c)), \\ M(u_a) \cup M(u_b) \cup \dots \cup M(u_c) = M(u) \cup M(u_1) \cup M(u_2) \cup \dots \cup M(u_m). \end{cases} \tag{9}$$

Then, we write

$$u \approx u_a + u_b + \dots + u_c. \tag{10}$$

$u$  is called a *reaction intermediate* generated from  $u_a, u_b, \dots, u_c$ .

**Remark 2.5.** If  $u$  is a reaction intermediate generated from  $u_a, u_b, \dots, u_c$ , then  $M(u)$  has the same contour as  $M(u_a) \cup M(u_b) \cup \dots \cup M(u_c)$ , but has internal

holes occupied by  $M(u_i)$  of loops  $u_1, u_2, \dots, u_m$ .

**Remark 2.6.**  $M(u_a) \cup M(u_b) \cup \dots \cup M(u_c)$  is called *strictly loop-transitive* if  $m = 0$ , i.e., there is a loop  $u$  such that  $M(u_a) \cup M(u_b) \cup \dots \cup M(u_c) = M(u)$ .

In nature, proteins interact with each other by forming “transient protein complexes”. In the loop model, interactions between loops are defined as fusions of loops.

**Definition 2.7. (Loop Interaction)** Loops  $u_a, u_b, \dots, u_c$  are called *interactable* if there is a reaction intermediate generated from  $u_a, u_b, \dots, u_c$ .

**Example 2.8.** In **Figure 3**,  $u_a$  and  $u_{ef}$  are interactable and their reaction intermediate is shown in the upper middle of the figure. On the other hand,  $u_{reg}$  and  $u_a$  are not interactable. For example, the “intermediate” shown in the lower middle of the figure encloses an “open” trajectory of length 6.

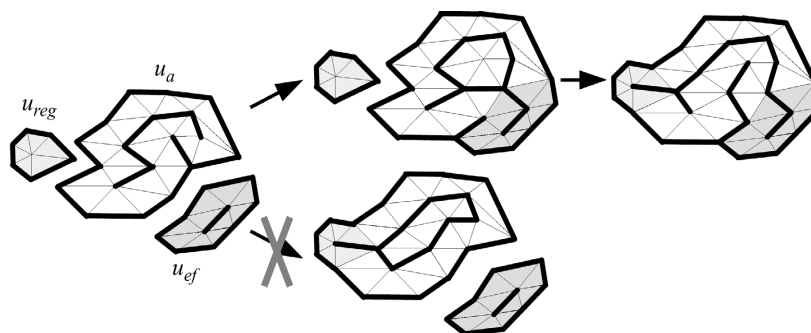
### 2.3. Allosteric Regulation in the Loop Model

“Allosteric regulation” is a type of interaction between two distal sites of a protein. For example, in the case of allosteric enzymes or receptors, the interaction of a protein with a molecule or protein (called “regulator”) at an “active” site is affected by the binding of another molecule (called “effector”) at a remote “allosteric” site. Allosteric sites have been attracted attention as targets for safer drug discovery [5]. However, allosteric site prediction still remains challenging [6]. Shown in **Figure 3** is a schematic diagram of allosteric regulation with three loops, where  $u_a$  interacts with “regulator”  $u_{reg}$  only after interaction with “effector”  $u_{ef}$ .

**Remark 2.9.** In the loop model, molecules are also represented as a loop of triangles or tetrahedra.

**Definition 2.10. (Allosteric Triplet)** Let  $u_a$ ,  $u_{reg}$ , and  $u_{ef}$  be loops of triangles. The triplet  $\langle u_a, u_{reg}, u_{ef} \rangle$  of loops are called an *allosteric triplet* if there are two reaction intermediates  $u_x, u_y$  such that

$$\begin{cases} u_x \approx u_a + u_{ef} \text{ and } u_y \approx u_{reg} + u_a + u_{ef}, \\ u \not\approx u_{reg} + u_a \text{ for any loop } u. \end{cases} \quad (11)$$



**Figure 3.** Allosteric triplet.  $u_a$  and  $u_{ef}$  generate a reaction intermediate  $u_x$  (upper middle).  $u_{reg}$  and  $u_a$  do not generate any reaction intermediate (lower middle).  $u_{reg}$  and  $u_x$  form a reaction intermediate  $u_y$  (right).

**Example 2.11.** Loops  $u_a, u_{reg}, u_{ef}$  in **Figure 3** form an allosteric triplet.

**Remark 2.12.** When effectors bind to proteins, they often change the conformation of the protein. In the loop model, since the loop model is a “topological” model, the conformational changes of proteins are considered to be absorbed into deformations of triangles in the mesh.

### 3. Flow Model of Proteins

In the flow model (**Figure 1(c)**), amino acid sequences are described as closed trajectories in a flow of triangles (2D model) or tetrahedra (3D model).

#### 3.1. Meshes

In this paper, a *triangular mesh* is a collection of triangles connected by a common edge. We write a triangular mesh  $M$  as a triplet of sets, *i.e.*,

$$M := \langle T, E(T), V(T) \rangle, \tag{12}$$

where

$$\begin{cases} T := \text{a set of triangles,} \\ E(T) := \text{the set of all edges of the triangles in } T, \\ V(T) := \text{the set of all vertices of the triangles in } T. \end{cases} \tag{13}$$

In this paper, we consider only the meshes such that (1) two triangles in  $T$  do not intersect except at an edge, and (2) the edges in  $E(T)$  are shared by at most two triangles in  $T$ . In particular,

$$E(T) = E_1(T) \cup E_2(T) \text{ (disjoint union)}, \tag{14}$$

Edges in  $E_1(T)$  are called the *boundary edges* of  $M$ . Binary set operations and binary relations between  $M_a$  and  $M_b$  are defined in the same way as for  $M(u_a)$  and  $M(u_b)$ .

**Definition 3.1. (Reaction Intermediate)** Let  $M = \langle T, E(T), V(T) \rangle$  be a triangular mesh.  $M$  is called *loop-transitive* if there are loops  $u, u_1, u_2, \dots, u_m$  such that

$$\begin{cases} T(u) \subset T, \\ E(T(u)) \supseteq E_1(T), \\ M = M(u) \cup M(u_1) \cup M(u_2) \cup \dots \cup M(u_m). \end{cases} \tag{15}$$

Then, we write

$$M \approx M(u). \tag{16}$$

$u$  is called a *reaction intermediate* generated from  $M$ .

**Remark 3.2.**  $M$  is called *strictly loop-transitive* if  $m = 0$ , *i.e.*, there is a loop  $u$  such that  $M = M(u)$ .

**Example 3.3.** Shown in the upper middle of **Figure 3** is a reaction intermediate generated from  $M(u_a) \cup M(u_{ef})$ . Shown in the right of **Figure 3** is a reaction intermediate generated from  $M(u_{reg}) \cup M(u_a) \cup M(u_{ef})$ .

In the same way, a *tetrahedral mesh* is a set of tetrahedra connected by a common face. We write a tetrahedra mesh  $M$  as a quartet of sets, *i.e.*,

$$M := \langle F, T(F), E(F), V(F) \rangle, \tag{17}$$

where  $F$  is a set of tetrahedra,  $T(F)$  is the set of all faces (*i.e.*, triangles) of the tetrahedra in  $F$ ,  $E(F)$  is the set of all edges of the tetrahedra in  $F$ , and  $V(F)$  is the set of all vertices of the tetrahedra in  $F$ .

### 3.2. Directed Elements of a Mesh

Let  $M = \langle T, E(T), V(T) \rangle$  be a triangular mesh. To specify flows on  $M$ , we consider directed elements of  $M$ , *i.e.*, directed vertices, directed edges, and directed triangles.

Since vertices have no direction, the set  $S_0(M)$  of all directed vertices in  $M$  is defined by

$$S_0(M) := V(T). \tag{18}$$

Edges in  $M$  are denoted by two endpoints, *i.e.*, the edge joining vertices  $v_a$  and  $v_b$  is denoted by  $v_a v_b$ . We make a distinction between two edges  $v_a v_b$  and  $v_b v_a$ , where  $v_a v_b$  is an edge with the direction from vertex  $v_a$  to vertex  $v_b$ . The corresponding edge in  $E(T)$  is denoted by  $|v_a v_b|$ , *i.e.*,  $|v_a v_b| = |v_b v_a| \in E(T)$ . Then, the set  $S_1(M)$  of all directed edges in  $M$  is defined by

$$S_1(M) := \{v_a v_b \mid |v_a v_b| \in E(T)\}. \tag{19}$$

Triangles in  $M$  are denoted by three vertices, *i.e.*, the triangle with three vertices  $v_a$ ,  $v_b$ , and  $v_c$  is denoted by  $v_a v_b v_c$ . We make a distinction between triangles with different order of vertices. For example,  $v_a v_b v_c \neq v_a v_c v_b$ , where  $v_a v_b v_c$  is a triangle with the ordered triplet  $\langle v_a, v_b, v_c \rangle$ . The corresponding triangle in  $T$  is denoted by  $|v_a v_b v_c|$ , *i.e.*,  $|v_a v_b v_c| = |v_a v_c v_b| \in T$ . Then, the set  $S_2(M)$  of all directed triangles in  $M$  is defined by

$$S_2(M) := \{v_a v_b v_c \mid |v_a v_b v_c| \in T\}. \tag{20}$$

### 3.3. Local Flows on a Mesh

Let  $M$  be a triangular mesh. A *local flow* on  $M$  is defined as a subset  $N$  of  $S_1(M)$ . Elements of  $N$  are called the *normal edges* of the local flow. We often denote a local flow  $N$  on  $M$  by  $\langle M, N \rangle$  (*i.e.*, as a pair of  $M$  and  $N$ ). Binary relations between  $\langle M_a, N_a \rangle$  and  $\langle M_b, N_b \rangle$  are defined by

$$\langle M_a, N_a \rangle \square \langle M_b, N_b \rangle \text{ if and only if } M_a \square M_b \text{ and } N_a \square N_b. \tag{21}$$

where  $\square$  stands for a binary relation such as  $\subset$ ,  $\supset$ , and others.

**Example 3.4.** Shown in **Figure 1(c)** is part of a local flow, where the normal edges are drawn with thick line segments.

**Remark 3.5.** Recall that a chain  $u$  of triangles can be recovered from the underlying mesh  $M(u)$  by giving the set  $N(u)$  of its normal edges (Lemma

2.2).

To perform “differential geometric analysis” of a local flow  $\langle M, N \rangle$ , we assign “gradient” to the edges in  $S_1(M)$ . First, we assign a “height” to the vertices in  $S_0(M)$ . Then, the “gradient” of the edges in  $S_1(M)$  is computed as the difference of the height along the edge. Finally, the normal edge of a triangle in  $S_2(M)$  is defined as the “steepest” edge of the triangles.

The *height*  $\alpha(v_a)$  of a vertex  $v_a \in S_0(M)$  is an integer-valued function defined on  $S_0(M)$ , i.e.,

$$\alpha : S_0(M) \rightarrow Z. \tag{22}$$

**Remark 3.6.** Vertices are considered to be lifted vertically from the mesh  $M$  through to the “height”.

The *gradient*  $\beta_\alpha(v_a v_b)$  of a directed edge  $v_a v_b \in S_1(M)$  with respect to  $\alpha$  is the difference in the height function  $\alpha$  along the edge, i.e.,

$$\beta_\alpha : S_1(M) \rightarrow Z, \beta_\alpha(v_a v_b) := \alpha(v_b) - \alpha(v_a) \tag{23}$$

Note that  $\beta_\alpha(v_a v_b) = -\beta_\alpha(v_b v_a)$ .

Let  $\beta$  be a  $Z$ -valued function on  $S_1(M)$ . The *normal edge*  $n_\beta(v_a v_b v_c)$  of  $v_a v_b v_c \in S_2(M)$  with respect to  $\beta$  is defined as the steepest (positive) edge of the triangle with respect to  $\beta$ , i.e.,  $n_\beta : S_2(M) \rightarrow S_1(M)$ ,

$$n_\beta(v_a v_b v_c) := \begin{cases} v_i v_j & \text{if } |\beta(v_i v_j)| > |\beta(v_j v_k)|, |\beta(v_k v_i)| \text{ and } \beta(v_i v_j) > 0 \\ v_j v_i & \text{if } |\beta(v_i v_j)| > |\beta(v_j v_k)|, |\beta(v_k v_i)| \text{ and } \beta(v_i v_j) < 0 \\ \emptyset & \text{otherwise} \end{cases} \tag{24}$$

where  $\{i, j, k\} = \{a, b, c\}$ .

**Remark 3.7.** We select edges with positive values as the normal edge of a triangle.

We denote the set of all normal edges of  $M$  with respect to  $\beta$  by  $N(\beta)$ , i.e.,

$$N(\beta) := \{n_\beta(v_a v_b v_c) \mid v_a v_b v_c \in S_2(M)\} \subset S_1(M). \tag{25}$$

The corresponding edges in  $E(T)$  are denoted by  $|N(\beta)|$ , i.e.,

$$|N(\beta)| := \{|n_\beta(v_a v_b v_c)| \mid v_a v_b v_c \in S_2(M)\} \subset E(T). \tag{26}$$

Triangles with one normal edge is called *regular*. Triangles with no normal edge and triangles with more than two normal edges are called *singular*. A local flow  $N$  is called *regular* if every triangle is regular. For example, a local flow corresponding to a loop (i.e.,  $|N(\beta)| = N(u)$ ) for some loop  $u$ ) is regular.

**Remark 3.8.** Triangles may have multiple normal edges because some edges are shared by two triangles.

A local flow  $\langle M, N \rangle$  is called *differentiable* if there is a  $Z$ -valued function  $\beta$  on  $S_1(M)$  such that  $N = N(\beta)$ .  $\beta$  is called the *vector field* of  $N$  and denoted by  $\beta(N)$ .

A differentiable local flow  $\langle M, N \rangle$  is called *integrable* if there is an  $Z$ -valued function  $\alpha$  on  $S_0(M)$  such that  $N = N(\beta_\alpha)$ .  $\alpha$  is called a *potential func-*

tion of  $N$  and denoted by  $\alpha(N)$ .

A differentiable local flow  $\langle M, N(\beta) \rangle$  is called 2-bounded if

$$0 < |\beta(v_a v_b)| \leq 2 \text{ for } \forall v_a v_b \in S_1(M). \tag{27}$$

**Proposition 3.9.** Let  $\langle M, N(\beta) \rangle$  be a 2-bounded differentiable local flow. Then,  $N(\beta)$  is regular if

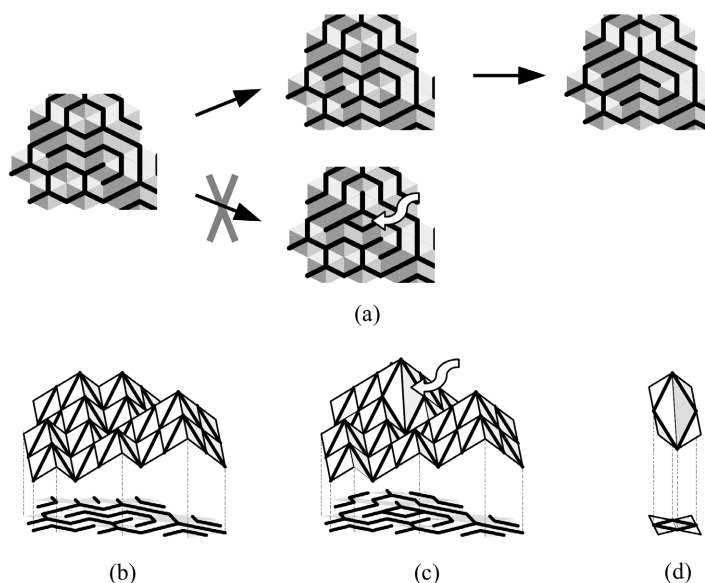
$$\beta(v_a v_b) + \beta(v_b v_c) + \beta(v_c v_a) = 0 \text{ for } \forall v_a v_b v_c \in S_2(M). \tag{28}$$

**Proof.** Let  $v_a v_b v_c \in S_2(M)$ . Since  $\langle M, N(\beta) \rangle$  is differentiable,  $\beta(v_i v_j) = 2$ ,  $\beta(v_i v_k) = 1$ , and  $\beta(v_k v_j) = 1$  for some  $i, j, k \in Z$  such that  $\{i, j, k\} = \{a, b, c\}$ . That is,  $n_\beta(v_a v_b v_c) = v_i v_j$ . On the other hand,  $v_i v_k$  and  $v_k v_j$  are not contained in  $N(\beta)$  because  $\beta(v_i v_k), \beta(v_k v_j) < 2$ . ■

**Remark 3.10.**  $\beta(v_a v_b) + \beta(v_b v_c) + \beta(v_c v_a)$  is called the *circulation* of  $\beta$  around a triangle  $v_a v_b v_c$ .

**Example 3.11.** By piling unit cubes up diagonally in the direction of  $(-1, -1, -1)$ , we obtain a regular local flow of triangles on the surface of the piled cubes (Figure 4). The normal edges are the vertical diagonals of the unit cubes. That is, each upper face of a unit cube is divided into two triangles by the vertical diagonal. Then, connecting triangles along the vertical diagonals, we obtain a flow on the surface of piled cubes. A triangular mesh  $M$  resides on the hyperplane  $x + y + z = 0$  and the height of a point  $p = (l, m, n) \in Z^3$  over

$$\pi(p) := \left( \frac{2l - m - n}{3}, \frac{-l + 2m - n}{3}, \frac{-l - m + 2n}{3} \right) \in M \tag{29}$$



**Figure 4.** Examples of local flows. (a) Top view of five local flows obtained by piling unit cubes. The normal edges are drawn in thick lines. If one more cube is put on the surface, the local flow will change as indicated by the arrows; (b) The regular flow shown in the upper middle of (a); (c) The non-regular flow shown in the lower middle of (a); (d) The singular triangles indicated by the S-shaped arrows in (a) and (c). Note that they are not obtained by dividing faces of a unit cube.

is given by

$$\alpha(\pi(p)) = l + m + n. \tag{30}$$

Note that the local flow of **Figure 4(c)** is differentiable but not 2-bounded due to the triangles pointed by the S-shaped arrow. On the other hand, the local flow of **Figure 4(b)** is differentiable and 2-bounded.

### 3.4. Interaction of Loops in a Flow

Let  $\langle M, N \rangle$  be a local flow. A *trajectory* of  $\langle M, N \rangle$  is a chain  $u$  of triangles such that

$$M(u) \subset M \text{ and } N(u) \subset N. \tag{31}$$

A closed trajectory of  $\langle M, N \rangle$  is called a *loop* of  $\langle M, N \rangle$ . A trajectory is called *maximal* if it cannot be extended further within  $M$ . Let  $\phi = \{u_a, u_b, \dots, u_c\}$  be a set of trajectories of  $\langle M, N \rangle$ .  $\phi$  is called a *flow* of  $\langle M, N \rangle$  if

$$M = M(u_a) \cup M(u_b) \cup \dots \cup M(u_c). \tag{32}$$

Let  $\langle M, N \rangle$  be a local flow, where  $M = \langle T, E(T), V(T) \rangle$ .  $\langle M, N \rangle$  is called *closed* if  $E_1(T) \subset N$ , i.e., the boundary edges of  $T$  are normal edges.  $\langle M, N \rangle$  is called *finite* if  $T$  consists of finite triangles.

**Proposition 3.12.** Let  $\langle M, N \rangle$  be a closed finite local flow. Then,

$$M = M(u_1) \cup M(u_2) \cup \dots \cup M(u_m) \tag{33}$$

for some loops  $u_1, u_2, \dots, u_m$  of  $\langle M, N \rangle$  if  $N$  is regular.

**Proof.** Since  $E_1(T) \subset N$ , trajectories of  $\langle M, N \rangle$  do not cross the boundary of  $M$ . Since  $N$  is regular, maximal trajectories of  $\langle M, N \rangle$  have no endpoint. The result follows immediately. ■

**Definition 3.13. (Loop Interaction)** Let  $u_a, u_b, \dots, u_c$  be loops of  $\langle M, N \rangle$ .  $u_a, u_b, \dots, u_c$  are called *interactable* if there is a reaction intermediate generated from  $u_a, u_b, \dots, u_c$ .

**Example 3.14.** Shown in the upper middle of **Figure 3** is a reaction intermediate generated from  $u_a, u_{ef}$ . Shown in the right of **Figure 3** is a reaction intermediate generated from  $u_{reg}, u_a, u_{ef}$ .

### 3.5. Allosteric Regulation in the Flow Model

**Definition 3.15. (Reaction Precursor)** Let  $M = \langle T, E(T), V(T) \rangle$  be a triangular mesh.  $M$  is called *pre-loop-transitive* if there is a loop  $u$  such that

$$\begin{cases} T(u) \subset T \\ E(T(u)) \supseteq E_1(T) \end{cases} \tag{34}$$

Then, we write

$$M \sim M(u). \tag{35}$$

$u$  is called a *reaction precursor* generated from  $M$ . By definition, reaction intermediates are reaction precursors.

**Remark 3.16.** If  $u$  is a reaction precursor of  $M$ , then  $M(u)$  may have internal holes with singular triangles inside.

**Example 3.17.** Shown in the lower middle of **Figure 3** is a reaction precursor generated from  $M(u_{reg}) \cup M(u_a)$ . It has a hole with two singular triangles, *i.e.*, the endpoints of the “open” trajectory of length 6.

**Proposition 3.18.** Let  $M$  be a triangular mesh. Suppose that there is a reaction precursor  $u$  generated from  $M$  of finite length, *i.e.*,  $M \sim M(u)$ . Let  $\langle M, N(\beta) \rangle$  be an 2-bounded differentiable local flow such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta) \rangle. \quad (36)$$

Then,  $u$  is a reaction intermediate (*i.e.*,  $M$  is loop-transitive) if  $\langle M, N(\beta) \rangle$  is integrable.

**Proof.** Since  $\langle M(u), N(u) \rangle$  is integrable, there is a potential function  $\alpha$  such that  $N(u) = N(\beta_\alpha)$ . Then,

$$\beta_\alpha(v_a v_b) + \beta_\alpha(v_b v_c) + \beta_\alpha(v_c v_a) = 0 \text{ for } \forall v_a v_b v_c \in S_2(M). \quad (37)$$

The result follows from Proposition 3.9 and Proposition 3.12 immediately. ■

**Definition 3.19. (Pre-Allosteric Triplet)** Let  $u_a$ ,  $u_{reg}$ , and  $u_{ef}$  be loops of  $\langle M, N \rangle$ . The triplet  $\langle u_a, u_{reg}, u_{ef} \rangle$  of loops are called a *pre-allosteric triplet* if there are two reaction precursors  $u_x, u_y$  such that

$$\begin{cases} M(u_a) \cup M(u_{ef}) \sim M(u_x) \text{ and } M(u_{reg}) \cup M(u_a) \cup M(u_{ef}) \sim M(u_y), \\ M(u_a) \cup M(u_{reg}) \not\sim M(u) \text{ for any loop } u. \end{cases} \quad (38)$$

**Corollary 3.20. (Conditions for loop interaction)** Let  $\langle M, N \rangle$  be a local flow. Let  $u_a$  and  $u_b$  be loops of  $\langle M, N \rangle$ . Suppose that there is a reaction precursor  $u$  generated from  $u_a$  and  $u_b$  of finite length, *i.e.*,

$$M(u_a) \cup M(u_b) \sim M(u). \quad (39)$$

Let  $\langle M, N(\beta) \rangle$  be an 2-bounded differentiable local flows such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta) \rangle. \quad (40)$$

Then,  $u_a$  and  $u_b$  are interactable if  $\langle M, N(\beta) \rangle$  is integrable.

**Corollary 3.21. (Conditions for allosteric triplet)** Let  $\langle u_a, u_{reg}, u_{ef} \rangle$  be a pre-allosteric triplet such that

$$M(u_a) \cup M(u_{ef}) \sim M(u_x) \text{ and } M(u_{reg}) \cup M(u_a) \cup M(u_{ef}) \sim M(u_y) \quad (41)$$

where  $u_x$  and  $u_y$  are reaction precursors of finite length. Let  $\langle M, N(\beta_x) \rangle$  and  $\langle M, N(\beta_y) \rangle$  be 2-bounded differentiable local flows such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta_x) \rangle \text{ and } \langle M(u), N(u) \rangle \subset \langle M, N(\beta_y) \rangle. \quad (42)$$

Then,  $\langle u_a, u_{reg}, u_{ef} \rangle$  is an allosteric triplet if  $\langle M, N(\beta_x) \rangle$  and  $\langle M, N(\beta_y) \rangle$  are integrable.

### 3.6. Post-Translational Modification in the Flow Model

“Post-translational modifications (PTMs)” are biochemical modifications of the

side chains of amino acids within a protein after their biosynthesis. They have a significant impact on the structure and function of proteins. For example, they play critical roles in regulating the stability of the 3D structure of proteins and their interactions with other molecules and proteins. In particular, the analysis of PTMs is important for the study of diseases, such as heart disease, cancer, and diabetes [7].

In the flow model, the effect of PTMs can be understood as a resolution of singularities of a local flow. That is, PTMs control the stability of a loop by inducing a resolution of the nearby singularity of the local flow.

It is helpful to consider the curvature of vertices to intuitively understand the geometric effects of singularity resolutions. Let  $M = \langle T, E(T), V(T) \rangle$  be a triangular mesh. Let  $v \in V(T)$ . The *curvature*  $K(v)$  of  $v$  is defined by

$$K(v) := k - 6, \tag{43}$$

where  $k$  is the number of edges incident on  $v$ .

**Remark 3.22.** In comparison to the continuous version of geometry,  $K$  corresponds to the ‘‘Gaussian curvature’’ of a plane. For example, a ‘‘saddle point’’ has a negative curvature, and a ‘‘point on a hemisphere’’ has a positive curvature.

**Example 3.23.** For triangular meshes obtained by piling unit cubes (Example 3.9.), the curvature is zero at all vertices (Figure 4).

**Example 3.24.** In Figure 1(c), a loop encloses two singular triangles. By splitting each of the singular triangles into two, we obtain a loop of length four (Figure 2(b) dark grey). The curvatures of the vertices of the singular triangles increase by 1 to become positive or zero (i.e., saddle-point). On the other hand, the new vertex obtained in the center has negative curvature  $-2$  (i.e., point on a hemisphere).

### 4. Cohomological Model of Proteins

In the cohomological model (Figure 1(c)), amino acid sequences are also described as closed trajectories in a flow of triangles (2D model) or tetrahedra (3D model) as in the case of the flow model. Here, we define ‘‘cohomology classes’’ of vector fields on a triangular mesh and rephrase conditions for ‘‘interaction’’ and ‘‘allosteric regulation’’ of loops using the language of cohomology.

#### 4.1. Functions on a Mesh

Let  $M$  be a triangular mesh. We denote the set of all ‘‘anti-symmetric’’ assignments of integers to the vertices, edges, and triangles in  $M$  by  $F_0(M)$ ,  $F_1(M)$ , and  $F_2(M)$ , respectively, i.e.,

$$F_0(M) := \{ \alpha : S_0(M) \rightarrow \mathbf{Z} \}, \tag{44}$$

$$F_1(M) := \{ \beta : S_1(M) \rightarrow \mathbf{Z} \mid \beta(v_a v_b) = -\beta(v_b v_a) \}, \tag{45}$$

$$F_2(M) := \{ \gamma : S_2(M) \rightarrow \mathbf{Z} \mid \gamma(v_a v_b v_c) = \gamma(v_b v_c v_a) = \gamma(v_c v_a v_b) = -\gamma(v_a v_c v_b) = -\gamma(v_c v_b v_a) = -\gamma(v_b v_a v_c) \}. \tag{46}$$

Elements of  $F_0(M)$  are called *scaler functions* (or *potential functions*) on  $M$ . Elements of  $F_1(M)$  are called *vector fields* on  $M$ .

**Remark 4.1.** In comparison to the continuous version of geometry,  $F_0(M)$  corresponds to “scaler functions (or potential functions)” on a space and  $F_1(M)$  corresponds to “vector fields” on a space.

Let  $X_1(M)$  be a subset of  $F_1(M)$  defined by

$$X_1(M) := \{ \beta \in F_1(M) : S_1(M) \rightarrow \{ \pm 1, \pm 2, \dots, \pm n \} \}, \quad (47)$$

where  $n = 2$  (if  $M$  is a triangular mesh) or  $n = 3$  (if  $M$  is a tetrahedral mesh). A vector field  $\beta \in F_1(M)$  is called *n-bounded* if  $\beta \in X_1(M)$ .

**Remark 4.2.** In comparison to the continuous version of geometry,  $X_1(M)$  corresponds to “differentiable vector fields” on a space.

### 4.2. Exterior Derivative Operator and Cohomology

Let  $M$  be a triangular mesh. Now let’s define “differentials” of functions on  $M$ .

*Discrete exterior derivative*  $d_i (i = 0, 1)$  is a mapping from  $F_i(M)$  to  $F_{i+1}(M)$  defined by

$$d_0 : F_0(M) \rightarrow F_1(M), \quad d_0 \alpha(v_a v_b) := \alpha(v_b) - \alpha(v_a), \quad (48)$$

$$d_1 : F_1(M) \rightarrow F_2(M), \quad d_1 \beta(v_a v_b v_c) := \beta(v_a v_b) + \beta(v_b v_c) + \beta(v_c v_a) \quad (49)$$

(Figure 5). Note that  $d_0 \alpha(v_a v_b) = -d_0 \alpha(v_b v_a)$  and  $d_1 \beta(v_a v_b v_c) = d_1 \beta(v_b v_c v_a) = d_1 \beta(v_c v_a v_b) = -d_1 \beta(v_a v_c v_b) = -d_1 \beta(v_c v_b v_a) = -d_1 \beta(v_b v_a v_c)$ .

**Remark 4.3.**  $d_0$  computes the difference of  $\alpha \in F_0(M)$  along an edge. On the other hand,  $d_1$  computes the circulation of  $\beta \in F_1(M)$  around a triangle.

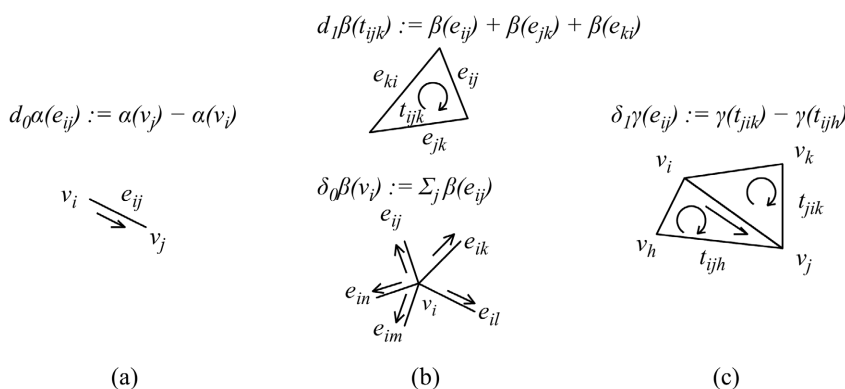
Let  $\beta \in F_1(M)$ .  $\beta$  is called *integrable* if  $\beta = d_0 \alpha$  for some  $\alpha \in F_0(M)$ .

**Lemma 4.4.**  $d_1 d_0 \alpha = 0$  for any  $\alpha \in F_0(M)$ .

**Proof.** It follows immediately from the definitions. ■

*Discrete exterior co-derivative*  $\delta_i (i = 0, 1)$  is a mapping from  $F_{i+1}(M)$  to  $F_i(M)$  defined by

$$\delta_0 : F_1(M) \rightarrow F_0(M), \quad \delta_0 \beta(v_a) := \sum_{\{v_b | v_a v_b \in S_1(M)\}} \beta(v_a v_b), \quad (50)$$



**Figure 5.** Discrete exterior derivative/co-derivative operators. In the figure,  $e_{ij}$  represents edge  $v_i v_j$  and  $t_{ijk}$  represents triangle  $v_i v_j v_k$ .

$$\delta_1 : F_2(M) \rightarrow F_1(M), \delta_1 \gamma(v_a v_b) := \gamma(v_b v_a v_c) - \gamma(v_a v_b v_d), \quad (51)$$

where  $v_b v_a v_c$  and  $v_a v_b v_d$  are the triangles that share the edge  $v_a v_b$  (Figure 5). Note that  $\delta_1 \gamma(v_a v_b) = -\delta_1 \gamma(v_b v_a)$ .

**Remark 4.5.**  $\delta_0$  computes the divergence of  $\beta \in F_1(M)$  at a vertex, i.e., the sum over the outbound arrows. On the other hand,  $\delta_1$  computes the difference of  $\gamma \in F_2(M)$  at a common edge.

**Example 4.6.** Figure 6(a) is a computation of  $d_1 d_0 \alpha = 0$  for  $\alpha$  on the left. Figure 6(b) is a computation of  $d_1(\beta_1)$  and  $\delta_1(\beta)$  for  $\beta$  in the center.

**Remark 4.7.** To learn more about discrete exterior derivative/co-derivative operators, see [8] or [9].

To define cohomology of vector fields on  $M$ , we consider a short sequence of sets given by

$$F_0 \xrightarrow{d_0} F_1 \xrightarrow{d_1} F_2. \quad (52)$$

Two subsets of  $F_1(M)$  are defined using the sequence, i.e.,

$$\text{Ker}(d_1) := \{\beta \in F_1(M) \mid d_1 \beta = 0\}, \quad (53)$$

$$\text{Im}(d_0) := \{d_0 \alpha \in F_1(M) \mid \alpha \in F_0(M)\}. \quad (54)$$

**Lemma 4.8.**  $\text{Im}(d_0) \subset \text{Ker}(d_1)$

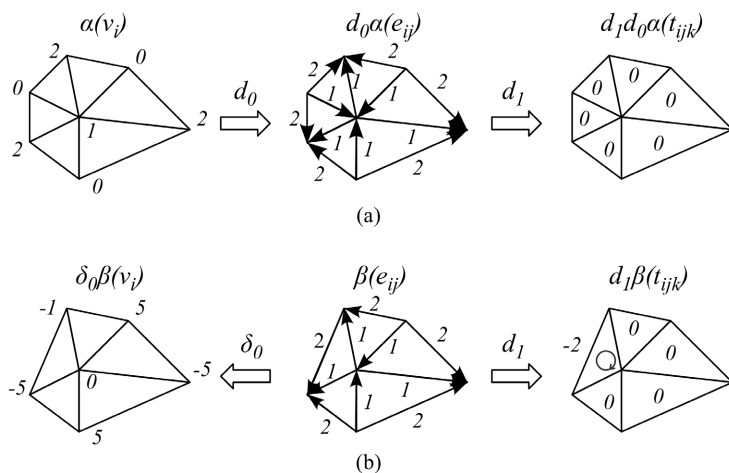
**Proof.** It follows immediately from Lemma 4.4. ■

Because of Lemma 4.8, we can consider the quotient set of  $\text{Ker}(d_1)$  by  $\text{Im}(d_0)$ .

**Definition 4.9. (Cohomology Class)** The (first) cohomology set  $H_1(M)$  is defined by

$$H_1(M) := \text{Ker}(d_1) / \text{Im}(d_0) \subset F_1(M) / \text{Im}(d_0). \quad (55)$$

Let  $\beta \in F_1(M)$ . The equivalence class  $\beta \bmod \text{Im}(d_0) \in H_1(M)$  is called the cohomology class of  $\beta$ .



**Figure 6.** Computation examples of discrete exterior derivative/co-derivative operators. In the figure,  $e_{ij}$  represents edge  $v_i v_j$  and  $t_{ijk}$  represents triangle  $v_i v_j v_k$ .

**Lemma 4.10.** If  $\beta \bmod \text{Im}(d_0) = 0 \in H_1(M)$ , then  $\beta = d_0\alpha$  for some  $\alpha \in F_0(M)$ .

**Proof.** It follows immediately from the definitions. ■

### 4.3. Cohomological Conditions for Allosteric Regulation

Let's rephrase some of the definitions given in Subsection 3.3. Let  $M$  be a triangular mesh. Then,

- 1) A local flow  $\langle M, N \rangle$  is *differentiable* if  $N = N(\beta)$  for some  $\beta \in F_1(M)$ ,
- 2) A differentiable local flow  $\langle M, N(\beta) \rangle$  is *integrable* if  $\beta$  is integrable (i.e.,  $\beta = d_0\alpha$  for some  $\alpha \in F_0(M)$ ),
- 3) A differentiable local flow  $\langle M, N(\beta) \rangle$  is *2-bounded* if  $\beta \in X_1(M)$ .

**Lemma 4.11.** Let  $\langle M, N(\beta) \rangle$  be a differentiable local flow. Then,  $d_1\beta = 0$  if  $\langle M, N(\beta) \rangle$  is integrable.

**Proof.** It follows immediately from Lemma 4.4. ■

Then, we obtain a cohomological description of conditions for loop interaction.

**Proposition 4.12.** Let  $M$  be a triangular mesh. Suppose that there is a reaction precursor  $u$  generated from  $M$  of finite length, i.e.,  $M \sim M(u)$ . Let  $\langle M, N(\beta) \rangle$  be an 2-bounded differentiable local flow such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta) \rangle. \tag{56}$$

Then,  $u$  is a reaction intermediate (i.e.,  $M$  is loop-transitive) if the cohomology class of  $\beta$  is zero.

**Proof.** It follows immediately from Proposition 3.18. ■

**Corollary 4.13. (Conditions for loop interaction)**

Let  $\langle M, N \rangle$  be a local flow. Let  $u_a$  and  $u_b$  be loops of  $\langle M, N \rangle$ . Suppose that there is a reaction precursor  $u$  generated from  $u_a$  and  $u_b$  of finite length, i.e.,

$$M(u_a) \cup M(u_b) \sim M(u). \tag{57}$$

Let  $\langle M, N(\beta) \rangle$  be an 2-bounded differentiable local flows such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta) \rangle. \tag{58}$$

Then,  $u_a$  and  $u_b$  are interactable if the cohomology class of  $\beta$  is zero.

**Corollary 4.14. (Conditions for allosteric triplet)**

Let  $\langle u_a, u_{reg}, u_{ef} \rangle$  be a pre-allosteric triplet such that

$$M(u_a) \cup M(u_{ef}) \sim M(u_x) \quad \text{and} \quad M(u_{reg}) \cup M(u_a) \cup M(u_{ef}) \sim M(u_y). \tag{59}$$

Let  $\langle M, N(\beta_x) \rangle$  and  $\langle M, N(\beta_y) \rangle$  be 2-bounded differentiable local flow such that

$$\langle M(u), N(u) \rangle \subset \langle M, N(\beta_x) \rangle \quad \text{and} \quad \langle M(u), N(u) \rangle \subset \langle M, N(\beta_y) \rangle. \tag{60}$$

where  $u_x$  and  $u_y$  are reaction precursors of finite length.

Then,  $\langle u_a, u_{reg}, u_{ef} \rangle$  is an allosteric triplet if the cohomology classes of  $\beta_x$  and  $\beta_y$  are zero.

## 5. Conclusions

We have considered protein interactions from the viewpoint of cohomology theory, using two-dimensional toy models of proteins. As a specific example, cohomological conditions for allosteric regulation are presented. In this paper, proteins are represented as loops of triangles and protein interactions are represented as fusions of loops. Then, cohomology classes of vector fields on proteins (*i.e.*, a triangular mesh) are defined using discrete exterior operators.

Cohomological conditions for loop-interaction (*i.e.*, protein interaction) are obtained as follows. First, we define reaction intermediates and their precursors generated from a given set of loops. By definition, loops interact if there is a reaction intermediate generated from the loops. Conditions for a precursor to be a reaction intermediate are then given using the language of “differential geometry”. That is, a precursor is a reaction intermediate if the local flow of the precursor is integrable. Finally, the cohomological conditions are obtained by rephrasing the differential geometric conditions using the language of “cohomology”. That is, a precursor is a reaction intermediate if the cohomology class of the vector field on the precursor is zero.

## 6. Discussion

Currently, since models of allosteric regulation only provide explanations of existing allostery, computer simulations are required to detect unknown allostery. However, the flow/cohomological model, despite its simplicity, is capable of explaining not only the existence of allostery but also its non-existence [10]. In particular, the model can predict the behavior of proteins: “if we remove the obstacle of allostery, we can obtain a new allosteric protein”.

In protein science, when considering protein-protein interactions, only local properties such as local shape complementarity are considered, mainly due to insufficient computer power. However, even local convexities on the surface of proteins are not formed locally, but as a result of global folding. One of the strengths of the flow/cohomological model is that we can consider both the shape of a protein and its folding structure at once. Then, the global properties of proteins can be described using the language of cohomology.

A drawback of the model is that it predicts nothing about the actual protein because it is a 2D model. Therefore, future research directions include the study of the 3D model, where a chains of tetrahedra would be a “backbone” (*i.e.*, streamline) of a flow, rather than a “cage” (*i.e.*, surface of a finite region) of a flow. Then, by detecting “turbulence in a flow” using the language of cohomology, we can predict the behavior of a protein.

Another direction is the study of surface flows, *i.e.*, triangular flows on the surface of folded chains of tetrahedra induced by a tetrahedral flow. As for the 2D model, the study of weaker conditions for loop interaction is also required.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Morikawa, N. (2017) Discrete Differential Geometry and the Structural Study of Protein Complexes. *Open Journal of Discrete Mathematics*, **7**, 148-164. <https://doi.org/10.4236/ojdm.2017.73014>
- [2] Morikawa, N. (2018) Global Geometrical Constraints on the Shape of Proteins and Their Influence on Allosteric Regulation. *Applied Mathematics*, **9**, 1116-1155. <https://doi.org/10.4236/am.2018.910076>
- [3] Liu, J. and Nussinov, R. (2016) Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Computational Biology*, **12**, e1004966. <https://doi.org/10.1371/journal.pcbi.1004966>
- [4] Pinčák, R., Kanjamapornkul, K. and Bartoš, E. (2020) Cohomology Theory for Biological Time Series. *Mathematical Methods in the Applied Sciences*, **43**, 552-579. <https://doi.org/10.1002/mma.5906>
- [5] Grover, A.K. (2013) Use of Allosteric Targets in the Discovery of Safer Drugs. *Medical Principles and Practice*, **22**, 418-426. <https://doi.org/10.1159/000350417>
- [6] Wu, N., Strömich, L. and Yaliraki, N.Y. (2022) Prediction of Allosteric Sites and Signaling: Insights from Benchmarking Datasets. *Patterns*, **3**, 100408. <https://doi.org/10.1016/j.patter.2021.100408>
- [7] Xu, H., Wang, Y., Lin, S., Deng, W., Peng, D., Cui, Q. and Xue, Y. (2018) PTMD: A Database of Human Disease-Associated Post-Translational Modifications. *Genomics, Proteomics & Bioinformatics*, **16**, 244-251. <https://doi.org/10.1016/j.gpb.2018.06.004>
- [8] Crane, C. (2021, Feb 24). Lecture 9: Discrete Exterior Calculus (Discrete Differential Geometry) [Video]. YouTube. <https://www.youtube.com/watch?v=-cUhuzwW-A>
- [9] Crane, C. (2022, May 2) Discrete Differential Geometry: An Applied Introduction. 1-172. <https://www.cs.cmu.edu/~kmc Crane/Projects/DDG/paper.pdf>
- [10] Morikawa, N. (2021) Two Mathematical Approaches to Inferring the Internal Structure of Proteins from Their Shape. *Global Journal of Science Frontier Research: F*, **21**, 1-25. <https://doi.org/10.34257/GJSFRFVOL21IS3PG1>