E(Q)AGNN-PPIS: Attention Enhanced Equivariant Graph Neural Network for Protein-Protein Interaction Site Prediction

Animesh, Rishi Suvvada, Plaban Kumar Bhowmick, Pralay Mitra

***Abstract*—Identifying protein binding sites, the specific regions on a protein's surface where interactions with other molecules occur, is crucial for understanding disease mechanisms and facilitating drug discovery. Although numerous computational techniques have been developed to identify protein binding sites, serving as a valuable screening tool that reduces the time and cost associated with conventional experimental approaches, achieving significant improvements in prediction accuracy remains a formidable challenge. Recent advancements in protein structure prediction, notably through tools like AlphaFold, have made vast numbers of 3-D protein structures available, presenting an opportunity to enhance binding site prediction methods. The availability of detailed 3-D structures has led to the development of Equivariant Graph Neural Networks (GNNs), which can analyze complex spatial relationships in protein structures while maintaining invariance to rotations and translations. However, current equivariant GNN methods still face limitations in fully exploiting the geometric features of protein structures. To address this, we introduce E(Q)AGNN-PPIS [1], an Equivariant Attention-Enhanced Graph Neural Network designed for predicting protein binding sites by leveraging 3-D protein structure. Our method augments the Equivariant GNN framework by integrating an attention mechanism. This attention component allows the model to focus on the most relevant structural features for binding site prediction, significantly enhancing its ability to capture complex spatial patterns and interactions within the protein structure. Our experimental findings underscore the enhanced performance of E(Q)AGNN-PPIS compared to current state-of-the-art approaches, exhibiting gains of 8.33% in the Area Under the Precision-Recall Curve (AUPRC) and 10% in the Matthews Correlation Coefficient (MCC) across benchmark datasets. Additionally, our method demonstrates robust generalization across proteins with varying sequence lengths, outperforming baseline methods.**

***Index Terms*—Graph neural networks (GNNs), Geometric GNNs, Proteins, Binding sites prediction.**

I. Introduction

PROTEINS, as the primary functional units in biological systems, are integral to the maintenance of structure and function in living organisms [1]. Comprehending their three-dimensional structure is critical for elucidating the underlying mechanisms of diverse biological processes, ranging from enzymatic catalysis, signal transduction and immune response to drug design. A critical challenge in protein structural modeling is the accurate prediction of protein binding sites, also referred to as protein-protein interaction (PPI) site prediction. This problem is of significant importance, as identifying protein binding sites facilitates a deeper understanding of their biological functions [2], subsequently enabling more efficient drug discovery and development processes [3]. Traditionally, protein structural modeling has relied on experimental techniques such as X-ray crystallography, cryo-electron microscopy, and nuclear magnetic resonance for determining protein complex structures [4]. However, these experimental biological methods are often characterized by high resource requirements, substantial time investments, and significant costs. These limitations have motivated the exploration of alternative approaches, particularly computational methods, which have recently demonstrated considerable promise in addressing these challenges.

Methods for PPI Site Prediction: Accurate binding site prediction necessitates the effective integration of physical, chemical, and geometric information. Recent advancements in structural biology, propelled by the merging of increased processing capacity and innovative learning techniques, have led to the development of various methods for protein-protein interaction prediction [5]–[8]. These learning based methods can be broadly categorized into two groups based on their approach to protein data processing: 1. Machine Learning (ML) and 2. Deep Learning (DL) based methods. ML-based methods typically use protein sequence and structural data, including raw sequences, evolutionary insights, and secondary structure information. These methods employ various ML algorithms, including the Naïve Bayes Classifier [9], Random Forest [5], Support Vector Machines (SVM) [10], and shallow neural networks [11], [12]. Despite the success of ML methods as majority in PPI site prediction, their inability to capture latent structural information often results in sub-optimal performance. The integration of DL methods, such as Convolutional Neural Networks (CNNs) [13], Recurrent Neural Networks (RNNs) [14], and Graph Neural Networks (GNNs) [15], has substantially enhanced the performance over traditional ML methods.

CNN and RNN Based Methods: CNN-based approaches [16]–[20] have been prominently employed to extract both local and global features from protein sequences and structures. For example, DeepPPIS [16] utilizes TextCNN to distill global features from protein sequences, whereas DeepSite [17] leverages a voxel-based 3-D CNN specifically for binding site prediction. Although CNN-based methods have demonstrated effectiveness in extracting features from proteins, they often

Manuscript created June, 2024; *(Corresponding Author: Animesh)*

Animesh, Rishi Suvvada and Plaban Kumar Bhowmick are with the Department of Artificial Intelligence, Indian Institute of Technology, Kharagpur (e-mails: animesh.sachan24794@kgpian.iitkgp.ac.in, rishisuvvada@gmail.com, plaban@cai.iitkgp.ac.in).

Pralay Mitra is with the Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur (e-mail: pralay@cse.iitkgp.ac.in).

[1]We will release the code upon the decision from reviewers.

overlook long-range dependencies in protein sequences. To overcome this limitation, DLPred [21] introduces an speedup version of the Long Short-Term Memory (LSTM) network, tailored to process sequence-based features derived from the Position-Specific Scoring Matrix (PSSM), alongside physical and physicochemical properties. Furthermore, DELPHI [18] and HN-PPISP [19] employ both Recurrent Neural Networks (RNNs) and CNNs to capture local and long-range features simultaneously. However, conventional convolutional networks face challenges in accurately identifying binding sites, potentially due to the irregularity of protein structures and the arbitrary rotation and translation of proteins in space [22]. Moreover, sequential input to Deep Learning methods is limited by the user-provided features and fails to fully exploit the hidden 3-D structure of proteins during information encoding. In response to these challenges, from the advent of AlphaFold [23], offering accurate complex protein structure predictions, has shifted focus towards methods utilizing tertiary structure information. Consequently, Graph Neural Network (GNN)-based methods have emerged as a robust solution, adept at capturing intricate structural details and enhancing binding site prediction accuracy.

Graph Neural Network Based Methods: GNN based methodologies are primarily divided into two categories based on their architecture and learning approaches: 1. Traditional GNNs and 2. Geometric GNNs. Traditional GNNs process graph-structured data by progressively refining node representations via a message-passing mechanism among adjacent nodes. For instance, GraphPPIS [6], utilizing a Graph Convolutional Neural Network (GCN) [24], frames the PPI binding site prediction as a node classification task and RGN [25] extends this by incorporating attention mechanism. This approach significantly surpasses earlier deep learning methods by accurately identifying interacting vs. non-interacting amino acid residues. Furthermore, AGAT-PPIS [7] extends the Graph Attention Network (GAT) [26] by incorporating edge features, thus strengthening structural awareness and enhancing translational and rotational invariance. Another noteworthy advancement is GHGPR-PPIS [8], it enhances the traditional GCN framework by integrating an edge-self attention mechanism and employing Heat Kernel techniques to amplify the efficacy of low-frequency filters. To date, GHGPR-PPIS stands as the most effective method for predicting PPI binding sites.

Limitations of Traditional GNN Methods and the Need for Equivariant Geometric GNNs: Despite the success of traditional GNN methods [6]–[8], [25] on PPI site prediction task, these methods often struggle to capture the intricate geometric dependencies present in protein structures, as they primarily rely on the graph topology and node features while neglecting the crucial spatial or geometric information. Traditional GNNs do not possess built-in equivariance properties in their architecture, making their predictions sensitive to rotations and translations of the input protein structures, which can lead to inconsistent results. Rotation and translation invariance properties are essential for proteins because their function is determined by their 3-D structure, which remains the same regardless of orientation or position in space. For instance, an enzyme's active site must be recognized irrespective of how the protein is oriented.

To overcome these limitations, equivariant Graph Neural Network (GNN) approaches, such as E(Q)GNN [27] and GVP-GNN [28], have been proposed. In equivariant Graph Neural Networks, 3-D coordinates are crucial as they capture the spatial relationships between atoms or residues. For example, in protein-protein docking, the precise 3-D arrangement of surface residues determines whether two proteins can bind, which cannot be inferred from sequence or graph topology alone [29]. These methodologies are constructed to be invariant under Euclidean group transformations, which encompass rotations, reflections, and translations, in addition to permutations. This characteristic renders them exceptionally apt for the modeling of molecular structures.

In this study, we aim to enhance the identification of binding sites through the application of geometric deep learning techniques. We advocate for the adoption of equivariant GNNs, with a specific focus on augmenting the GVP-GNN architecture [28] through the integration of an attention mechanism. This approach exploits the capabilities of equivariant message passing while handling both scalar and vector features, thus enabling the detailed capture of geometric information present in protein structures. The main contributions of our work are outlined as follows:

- We introduce E(Q)AGNN-PPIS, an equivariant geometric graph neural network architecture that leverages geometric information, designed with a focus on PPI site prediction but potentially applicable to other structural graphs as well. E(Q)AGNN-PPIS is the first method to leverage the expressive power of equivariant message passing, incorporating both scalar and vector features, while introducing an attention mechanism to selectively focus on the most relevant features and interactions during message passing in PPI site prediction task.
- We demonstrate the superior performance of E(Q)AGNN-PPIS on the PPI site prediction task, surpassing state-of-the-art methods. Our model achieves significant improvements in terms of precision (5.45%), recall (9.68%), auprc (8.33%) and F1-score (6.90%), highlighting its effectiveness in accurately identifying binding sites.
- We conduct extensive ablation studies and present visualizations to gain insights into the inner workings of E(Q)AGNN-PPIS. The findings reveal the importance of the attention mechanism and the equivariant message passing in capturing the geometric intricacies of protein structures.

The rest of the manuscript is structured as follows: Section II introduces the fundamental concepts that inspire our method and datasets for training and evaluation. Section III presents E(Q)AGNN-PPIS, our proposed method, and evaluation strategies. Section IV details about experimental setup with results, discussion and ablation studies, followed by conclusion in section V.

II. Materials and Definitions

In this section, we present a comprehensive overview of the dataset employed in our study, along with a detailed description

of the data representation tailored to the problem at hand. To facilitate a clear understanding of the problem and its intricacies, we also provide the necessary definitions and notations.

A. Datasets Utilized

In order to guarantee an equitable comparison and thorough assessment of the proposed approach, we employ the widely-used benchmark dataset introduced in the state-of-the-art study GHGPR-PPIS [8]. This dataset, which is derived originally from GraphPPIS [6] and refined further, has been extensively utilized in various studies and serves as a standard for evaluating the performance of PPI site prediction methods. The dataset consists of many subsets: a training set labelled as Train_335-1, and three test sets named Test_315-28, Test_60-0, and Ubtest_31-6. Every subset in the dataset is skewed towards negative samples (non-binding sites), making the classification problem challenging. The Train_335-1 and Test_60-0 subsets were principally utilised as benchmark datasets for training and evaluating the proposed model in this work. The other two subsets, Test_315-28 and Ubtest_31-6, were employed to evaluate the model's ability to generalize to new, unseen data. The details of the datasets used in this study is provided in table I. The extensive details about the datasets can be found in the supplementary dataset section.

TABLE I
DESCRIPTIVE STATISTICS OF DATASETS USED IN THIS STUDY DERIVED FROM GHGPR-PPIS [8].

Dataset	Protein Complexes	BS[a] Count	Non-BS[b] Count	Excluded[c] Chains	Imbalance[d] Ratio
Train_335-1	209	10,336	58,872	1	**1:6**
Test_60-0	58	2,075	11,069	0	**1:5**
Test_315-28	217	8,566	51,810	28	**1:6**
Ubtest_31-6	25	711	5,206	6	**1:7**

[a] **BS**: Binding Sites, where amino acids represents the binding sites. **BS Count**: Total number of true binding sites present in the dataset.
[b] **Non-BS Count**: Total number of non-binding sites present in the dataset.
[c] **Excluded Chains**: Number of removed chains that doesn't align with their corresponding structure from original dataset [6].
[d] **Imbalance Ratio**: Ratio of binding sites and non-binding sites present in the dataset.

B. Geometric Graph Construction and Input features

1) Protein Geometric Graph Representation: In this work, we represent each protein structure input, consisting of sequence of n amino acids, as undirected graph $G_p = (V, E, A, \vec{X})$. Each graph incorporates both scalar and vector features to describe the 3-D structure of proteins. The node feature matrix V has rows $v_i \in \mathbb{R}^{D_v}$ for each of the N_v nodes, where v_i represents the feature vector of node i. The edge feature matrix E contains vectors $e_{ij} \in \mathbb{R}^{D_e}$ for each edge between nodes i and j, where an edge exists if $A_{ij} = 1$ in the adjacency matrix $A \in \mathbb{R}^{N_v \times N_v}$ of graph G_p. $\vec{X} = \{\vec{x}_i\}_{i=1,\ldots,N_v}$ are the positions of central node atoms C_α of amino acids in 3-D Euclidean space, and $\vec{x}_i \in \mathbb{R}^3$.

Each amino acid in the protein sequence is represented as a node in the graph G_p. An amino acid is made up of four *backbone atoms*[2] and a set of side-chain atoms, which are located in 3-D Euclidean space. The construction of each edge e_{ij} is based on a threshold Euclidean distance between the central carbon atom C_α of the corresponding amino acids represented as node i and j. The 3-D coordinates of the C_α atom are extracted from the protein's PDB file. The Euclidean distance d is computed as $d_{ij} = \|\vec{x}_i - \vec{x}_j\|_2$ for every pair of C_α atoms in the protein chain's amino acids. By this we represent a protein chain as an adjacency matrix A, where:

$$A_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq \theta \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Here θ denotes distance threshold, considered as 14Å (angstroms) following the works [6]–[8]. This approach explicitly enforces the structural locality as an inductive bias because nodes farther apart generally does not participate in the interaction. However, stacking multiple GNN layers enables capturing long range interactions as well. We also capture the effect of varying distance threshold as a part of ablation study IV-D.

This study focuses on the utilization of scalar node and edge features which are consistent with the state-of-the-art method GHGPR-PPIS [8]. However we also introduce edge vector features which will be discussed in this section.

2) Node Features: The node features of the geometric protein graph integrate both sequential and structural information of protein chains, offering a comprehensive 62-dimensional feature vector for each amino acid residue. Sequence-based features are derived from Position-Specific Scoring Matrices (PSSM) and Hidden Markov Model (HMM) profiles. PSSM profiles are generated utilizing PSI-BLAST [30], which is used to query the sequence against the UniRef90 database [31], performing three iterations with an E-value of 0.001, while HMM profiles are constructed via HHblits [32], each yielding a feature matrix with dimensions $N_v \times 20$, where N_v represents the length of protein sequence. These matrices are normalized to ensure uniform data scaling across various input sets. Structural features are sourced from the Dictionary of Secondary Structure of Proteins (DSSP) [33], providing a 14-dimensional secondary structural vector per residue. This includes a nine-dimensional[3] one-hot encoded matrix representing secondary structure categories. An additional four dimensions represent the torsion angles ϕ and ψ, transformed into sine and cosine values, enhancing rotational invariance. The last dimension quantifies relative solvent accessibility, computed from solvent accessible surface areas. Each amino acid residue is characterised by 7 atomic features, which include atomic mass, B-factor, atom type (side chain or backbone), electrostatic charge, number of bonded hydrogen atoms, presence of aromatic residue, and Van der Waals radius of the atom. These features provide detailed information at

[2]The atoms C_α, C, N and O. The alpha carbon (C_α) is the central carbon atom in every amino acid residue.

[3]H and G represent the alpha-helix and 3_{10} helix, respectively. I: -helix and E represents an extended beta-strand. B: beta-strand, T: turns, S: bend, The variable C represents secondary structures of proteins that are not classified into any specific category, as well as those that are unknown.

the atomic level, which is essential for accurately modelling protein interactions. Finally, the pseudo-position embedding feature (PPEF) encodes the relative positional information of each amino acid residue, benchmarked against a reference residue which we took as the first residue of the protein chain, captures spatial relationships critical for understanding residue interactions.

3) Edge Features: Each edge e_{ij} in the set of edges, where $i \neq j$, represents a connection between two amino acid residues in the graph and have the following associated features:

- To encode the direction, we calculate the unit vector in the direction from $\vec{x}_i$ to $\vec{x}_j$ by normalizing the vector $\vec{x}_i - \vec{x}_j$. This unit vector serves as a directional feature for the edge connecting center C_α atoms for node i and j.
- We encode the Euclidean distance between the residues $\|\vec{x}_i - \vec{x}_j\|_2$ in terms of Gaussian radial basis function.[4]
- The cosine of the angle formed by the vector connecting the residues and a reference vector pointing towards the first amino acid residue of the protein chain.

We summarize all the features used in this study in table II with their corresponding dimensions.

TABLE II
SUMMARY OF NODE AND EDGE FEATURES USED IN THE E(Q)AGNN-PPIS MODEL AND THEIR DIMENSIONS.

Feature Type	**Dimension**
PSSM	$N_v \times 20$
HMM	$N_v \times 20$
Atomic Features	7
DSSP	14
Pseudo-Position Embedding Feature (PPEF)	1
Euclidean Distance (Edge) as Gaussian RBFs	16
Cosine of Angle (Edge)	1
Unit Vector to encode direction (Edge)	3

Note: Node and Edge features are separated by a horizontal line.

C. Message Passing Neural Networks (MPNNs)

Message Passing Neural Networks (MPNNs) [34] extend the concept of Graph Neural Networks (GNNs) [35], [36] by providing a more general framework for learning representations of graphs. MPNNs define a parameterized function that maps graph structures to feature vectors, operating at both node and graph levels. This approach enables more flexible and expressive modeling of graph-structured data. Since MPNNs [34] employ shared learnable layers across all nodes in the graph, they maintain permutation equivariance, ensuring that the output remains consistent under any reordering of the input nodes.

However, when dealing with geometric graphs, such as those representing protein structures, additional challenges arise. Standard GNNs directly incorporating atomic 3-D coordinates as scalar features would disrupt the equivariance under Euclidean transformations of the input system. Consequently, the primary challenge for Geometric GNNs is to develop an effective method to encode and process geometric graphs while preserving geometric symmetries such as translation invariance, rotational equivariance and reflection equivariance. At a higher abstraction level, the transformation of scalar and vector features from layer t to $t+1$ is facilitated through learnable message and update functions, denoted as MSG and UPD respectively, coupled with a fixed permutation-invariant operator $\oplus$ (such as mean or sum):

$$m_{ij}^{(t)}, \vec{m}_{ij}^{(t)} := MSG\left(s_i^{(t)}, s_j^{(t)}, \vec{v}_i^{(t)}, \vec{v}_j^{(t)}, \vec{x}_{ij}\right) \tag{2}$$

$$s_i^{(t+1)}, \vec{v}_i^{(t+1)} := UPD\left(\left(s_i^{(t)}, \vec{v}_i^{(t)}\right), \bigoplus_{j \in N_i}\left(m_{ij}^{(t)}, \vec{m}_{ij}^{(t)}\right)\right) \tag{3}$$

with scalar messages $m_{ij}^{(t)}$ and vector messages $\vec{m}_{ij}^{(t)}$, $s^{(t)}$ denotes scalar features, $\vec{v}^{(t)}$ denotes vector features representing geometric features such as 3-D positional vector, and $\vec{x}_{ij} = \vec{x}_i - \vec{x}_j$ denoting relative position vector.

D. Invariance and Equivariance

To fully understand the importance of geometric symmetries in MPNNs and Geometric GNNs, it's crucial to define two key properties: invariance and equivariance. These properties are fundamental to designing models that can effectively handle geometric transformations of input data.

Definition II.1 (Invariance). *Let $f : \mathcal{G} \to \mathcal{Y}$ be a function that maps a graph $G \in \mathcal{G}$ to an output $Y \in \mathcal{Y}$. The function f is said to be invariant to a transformation $T : \mathcal{G} \to \mathcal{G}$ if for any graph $G \in \mathcal{G}$, $f(T(G)) = f(G)$. In other words, the output of the function remains unchanged under the transformation T.*

Definition II.2 (Equivariance). *Let $f : \mathcal{G} \to \mathcal{Y}$ be a function that maps a graph $G \in \mathcal{G}$ to an output $Y \in \mathcal{Y}$, and let $T_\mathcal{G} : \mathcal{G} \to \mathcal{G}$ and $T_\mathcal{Y} : \mathcal{Y} \to \mathcal{Y}$ be transformations on the input and output spaces, respectively. The function f is said to be equivariant to transformations $T_\mathcal{G}$ and $T_\mathcal{Y}$ if for any graph $G \in \mathcal{G}$, $f(T_\mathcal{G}(G)) = T_\mathcal{Y}(f(G))$. In other words, the transformation of the output is equivalent to applying the function to the transformed input.*

In the context of geometric graphs, these properties are particularly important. Let $G_p = (V, E, \vec{X})$ be a geometric graph with $\vec{X}$ being 3-D geometric vectors (positions, velocities, etc) that are steerable i.e. they transform predictably under the actions of the $E(3)$ group (rotations/translations/reflections), and h being non-steerable features of nodes V. For a geometric graph neural network function f to be $E(3)$-equivariant, it must satisfy: $f(T \cdot \vec{X}, h) = T \cdot f(\vec{X}, h), \forall \vec{X} \in \mathbb{R}^{3\times m}, h \in \mathbb{R}^{D_v}, T \in E(3)$. Similarly, f is invariant if, $f(T \cdot \vec{X}, h) = f(\vec{X}, h)$. The group action $\cdot$ is instantiated as $T \cdot \vec{X} := \vec{X} + b$ for translation $b \in \mathbb{R}^3$ and $T \cdot \vec{X}^{(t)} := O\vec{X}$ for rotation/reflection $O \in \mathbb{R}^{3\times 3}$.

E. Problem Statement

Let's define the binding criterion $C(v_i)$ for a node v_i to be classified as binding site or not in a geometric protein graph G_p as:

[4]We employ 16 Gaussian radial basis functions (RBFs) with centers uniformly distributed between 0 and 20 angstroms to effectively capture spatial relationships.

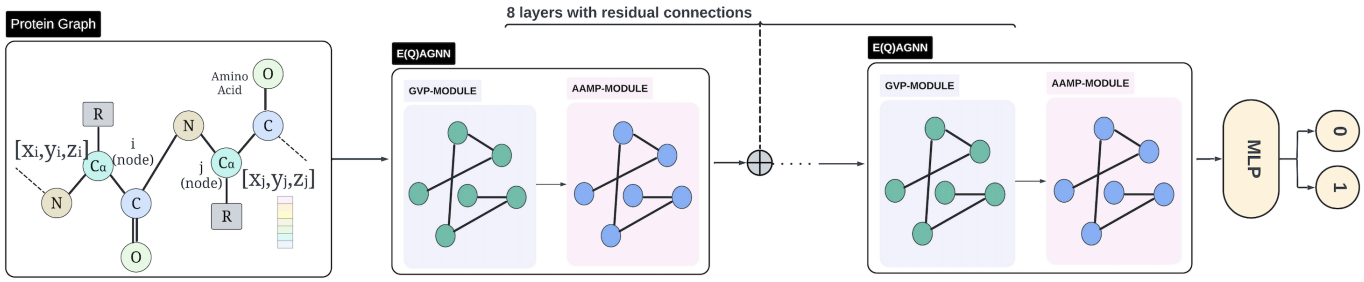

Fig. 1. End-to-end architecture of E(Q)AGNN-PPIS for protein-protein interaction site prediction. The model processes a protein graph representation depicting amino acids as nodes with their constituent atoms and 3-D coordinates, through 8 E(Q)AGNN layers, each containing Geometric Vector Perceptron (GVP) and Attention Aggregated Message Passing (AAMP) modules. These layers are connected via residual connections. The final node embeddings are fed into a Multi-Layer Perceptron (MLP) for binary classification of each amino acid as a binding site (1) or non-binding site (0).

$$C(v_i) = \begin{cases} 1 & \text{if } v_i \text{ is part of a binding site} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Given this binding criterion and a geometric protein graph G_p, our objective is to develop and train an Equivariant Graph Neural Network (GNN) model $f(G_p, C)$ to ensure consistent predictions regardless of the protein's orientation in 3-D space, which is crucial for accurate binding site identification. This model aims to predict whether any given node v_i in the protein graph G_p is a potential binding site or not. The challenge lies in leveraging the structural and topological information encoded in G_p to effectively classify the nodes, thus enhancing our ability to predict interaction sites in proteins with high accuracy.

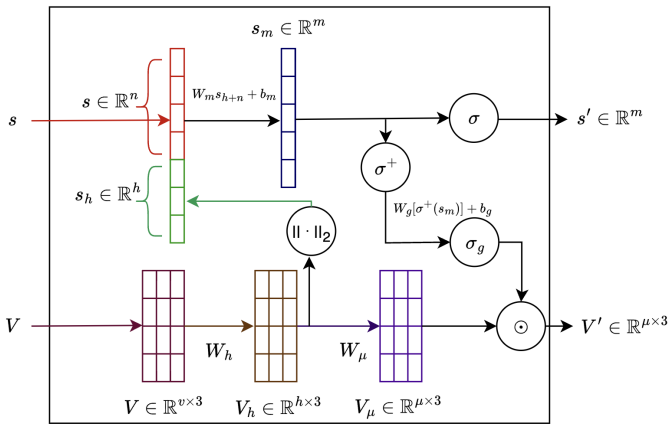

Fig. 2. Schematic representation of GVP-Module [37].

III. Methodology and Evaluation

In this section, we present our proposed method, including it's mathematical formulation along with evaluation strategies.

A. E(Q)AGNN-PPIS

1) GVP Embeddings: Geometric Vector Perceptron (GVP-Module) [28] is a simple module to learn scalar and vector valued functions over geometric vectors and scalars. In this work, we employ an enhanced version of the GVP-Module incorporating vector gating [37], which addresses limitations in the original formulation, particularly for atomic-level structure graphs, where individual atoms may not inherently possess orientation information (vector features) at the input level. Vector gating allows the propagation of information from scalar channels into vector channels. The GVP-Module with vector gating [37] as shown in Figure 2, which takes the scalar $s \in \mathbb{R}^n$ and vector $V \in \mathbb{R}^{v\times 3}$ features as a tuple $(s; V)$ and returns the new scalar and vector features tuple $(s'; V') \in \mathbb{R}^m \times \mathbb{R}^{\mu\times 3}$. In Figure 2 W_h, W_μ, W_m are trainable weight matrices specific to the GVP-Module, b_m, b_g are bias vectors, σ is a non-linear activation function, and $\odot$ denotes element-wise multiplication.

The GVP-Module respects crucial properties for geometric data processing: invariance, equivariance, and expressiveness. It preserves equivariance and invariance under 3-D Euclidean transformations T, such that $GVP(s, T(V)) = (s', T(V'))$ when $GVP(s, V) = (s', V')$, as validated in GVP-GNN [28]. The vector gating mechanism enhances the module with universal approximation capability for rotation and reflection equivariant functions $T : \mathbb{R}^{v\times 3} \to \mathbb{R}^3$ [37].

2) Attention Aggregate Message Passing - AAMP Module: In the proposed E(Q)AGNN-PPIS architecture, the GVP-Module processes both node and edge attributes of the protein graph while preserving equivariance. This processed information is then used in an attention-based mechanism to aggregate data from neighboring nodes. The proposed message passing scheme formulates messages by combining both node and edge features, as detailed in Equation 5. During the aggregation process, the architecture assigns attention scores to these messages from neighboring nodes. These scores effectively weight the contribution of each neighbor, allowing the model to focus on the most relevant information when updating a node's representation. This approach enables the model to capture complex spatial relationships and interactions within the protein structure, which is essential for accurate prediction of protein-protein interaction sites.

$$m_{ij} = \mathcal{S}\left(\mathfrak{h}_{\mathfrak{v}}^{i} || \mathfrak{h}_{\mathfrak{e}}^{j\to i}\right) \tag{5}$$

Central to the E(Q)AGNN-PPIS model is the attention mechanism, pivotal for determining the weights of the messages exchanged between nodes in the graph. This mechanism calculates attention scores α_{ij}, which measure the importance of the messages being transmitted from node j to node i. The

adjacency information A_{ij} is utilized to modulate the attention scores, ensuring that only connected nodes, as defined by the adjacency matrix in Equation 1, contribute to the updated node embeddings. This approach enables the model to focus on the most informative interactions while ignoring irrelevant ones. To calculate the attention scores, we utilize scalar node features, specifically protein node features, as demonstrated in Equations 6 and 7. This design decision enables the model to effectively reflect the intrinsic characteristics of the protein nodes and their impact on the attention mechanism. Once the attention-modulated messages are computed, they are aggregated for each node, effectively combining the weighted contributions from all neighboring nodes.

$$\beta_{ij} = \text{LeakyReLU}\left(a^T\left(W^{\alpha}[W^v h_v^i || W^v h_v^j]\right)\right) \tag{6}$$

$$\alpha_{ij} = \text{softmax}\left(\beta_{ij} \cdot A_{ij}\right) \tag{7}$$

Following the message aggregation step, the node features are updated using Layer Normalization, as shown in Equation 8. This normalization technique is employed to stabilize the learning process by reducing the internal co-variate shift.

$$\mathfrak{h}_{\mathfrak{v}}^i \leftarrow \text{LayerNorm}\left(\mathfrak{h}_{\mathfrak{v}}^i + \frac{1}{k_i}\text{Dropout}\left(\sum_{j:j\neq i} \alpha_{ij} m_{ij}\right)\right) \tag{8}$$

Here, in Equation 5, $\mathcal{S}$ represents a sequence of three GVP-Module layers. $\mathfrak{h}_{\mathfrak{v}}^i = (h_v^i \| \vec{h_v}^i)$ and $\mathfrak{h}_{\mathfrak{e}}^{j\to i} = (h_e^{j\to i} \| \vec{h_e}^{j\to i})$ are the embeddings of the node i and edge $j \to i$ having both scalar and vector features. $W^{\alpha}, W^v$and a^T are learnable parameters metrics of liner layer at different locations and $\|$ denotes the concatenation operation. A_{ij} represents the elements from the adjacency matrix as defined in Equation 1; LeakyReLU serves as the non-linear activation function, characterized by a negative slope of 0.2; k_i denotes the count of incoming messages for node i. The LayerNorm [38] is the normalization function applied to each node feature vector independently, ensuring that each feature is normalized based on its own statistics.

Between the propagation steps in the graph, a pointwise feed-forward layer, as shown in equation 9, is employed to update the embeddings of all nodes i. This additional layer enables the model to capture more complex interactions between the nodes.

$$\mathfrak{h}_{\mathfrak{v}}^i \leftarrow \text{LayerNorm}\left(\mathfrak{h}_{\mathfrak{v}}^i + \text{Dropout}\left(\mathcal{S}\left(\mathfrak{h}_{\mathfrak{v}}^i\right)\right)\right) \tag{9}$$

Here $\mathcal{S}$ at Equation 9 in the sequence of 2 GVP-modules. The graph propagation and pointwise feed-forward layers update not only the scalar but also the vector features for each node in the graph. Our end-to-end architecture, illustrated in Figure 1, employs 8 layers of E(Q)AGNN interconnected via residual connections. This design mitigates the oversmoothing issue while enhancing the capture of structural information.

3) Building Equivariant Features: In protein structure analysis, we leverage rich atomic-level information to construct equivariant features, distinguishing our approach from typical point cloud datasets. We compute both node and edge vector features, which satisfy specific equivariance conditions (for proof, refer to Appendix C-A and C-B), ensuring our model maintains geometric consistency under spatial transformations. For nodes, we compute vector features $\vec{h_v^i}$ as described in Proposition 1. For edges, we introduce vector features $\vec{h_e^{j\to i}}$ that capture directional information, as detailed in Proposition 2. These vector features are complemented by scalar features for both nodes and edges, derived from various physicochemical properties and sequence-based information (see Sections II-B2 and II-B3). This hybrid approach allows our E(Q)AGNN-PPIS model to simultaneously process both structural and biochemical aspects of protein interactions while maintaining appropriate equivariance properties.

Proposition 1. *Let $\vec{x_i} \in \mathbb{R}^3$ be the position vector of the ith node representing an amino acid, and $\mathcal{N}_i$ be the set of neighbors of node i. The node vector feature $\vec{h_v}^i$ defined as:*

$$\vec{h_v}^i = \frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i}(\vec{x_j} - \vec{x_i}) \tag{10}$$

is invariant under translations and equivariant under the orthogonal group O(3), which includes rotations and reflections.

Proposition 2. *Let $\vec{x_i}, \vec{x_j} \in \mathbb{R}^3$ be the position vectors of nodes i and j. The edge vector feature $\vec{h_e}^{j\to i}$ defined as:*

$$\vec{h_e}^{j\to i} = \frac{\vec{x_j} - \vec{x_j}}{\|\vec{x_j} - \vec{x_j}\|} \tag{11}$$

is invariant under translations and equivariant under the orthogonal group O(3), which includes rotations and reflections.

4) Model Equivariance: Having introduced all components of our E(Q)AGNN-PPIS model, including the GVP Embeddings, Attention Aggregate Message Passing (AAMP) Module, and the construction of equivariant features, we now establish the equivariance properties of the entire model. The The proposition 3 shows that E(Q)AGNN-PPIS preserves the equivariance properties of its input features throughout its operations.

Proposition 3. *The E(Q)AGNN-PPIS model, including its attention mechanism defined in Equations (5)–(9), is equivariant under the actions of the orthogonal group $O(3)$.*

Proof. See Appendix C-C □

B. Performance Evaluation Metrics

To evaluate our E(Q)AGNN-PPIS method for predicting protein-protein interaction (PPI) binding sites the metrics like Accuracy, Precision, Recall, F1 Score, Matthews Correlation Coefficient (MCC), and Area Under the Precision-Recall Curve (AUPRC) have been used in this study. This comprehensive approach allows for direct comparison with state-of-the-art methods in PPI site prediction. As evident from Table I, the datasets used in this study exhibit a significant class imbalance, which is inherent to the PPI binding site prediction problem. Consequently, the most informative and commonly used metrics for evaluating model performance in such scenarios are F1 Score, MCC, and AUPRC, as they provide a more comprehensive assessment of the model's ability to handle

class imbalance effectively. The equations to compute these metrics are given in Appendix B.

IV. Experiments and Results

A. Implementation Details and Hyper-Parameters

For the development of the E(Q)AGNN-PPIS, we employed the PyTorch v2.2.1[39] and PyTorch-Geometric v2.4.0 [40] libraries. Optimization was performed using the Adam [41] optimizer over 100 epochs with an initial learning rate of 0.0005, targeting the minimization of the Binary Cross-Entropy loss as specified in equation 12. The architecture included eight layers of E(Q)AGNN, incorporating residual connections to counter over-smoothing effects [42]. To reduce overfitting, we applied a dropout rate of 0.1 throughout the training process. A plateau learning rate scheduler was also utilized, adjusting the learning rate based on stagnation in the validation loss. To compensate for class imbalances in the dataset, we adjusted the class weights in the Binary Cross-Entropy loss function to a 1:4 ratio, with the positive class (interaction sites) weighted more heavily than the negative class (non-interaction sites).

$$BCE = -\frac{1}{n}\sum_{i=1}^{n}\left(y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right) \quad (12)$$

where n is the number of samples, and y_i and $\hat{y}_i$ represent the actual and predicted labels, respectively.

A detailed hyper-parameters search and finally used hyper-parameters for training are detailed in table III.

TABLE III
Hyper-Parameter search for E(Q)AGNN-PPIS.

H. Parameter[a]	Value	H. Parameter[a]	Value
Epochs	{50, **100**, 150}	Learning Rate	{0.01, 0.001, **0.0005**}
Dropout	{0.2, 0.15, **0.1**}	E(Q)AGNN Layers	{4, 6, **8**}

Used hyper-parameters for training are in **bold**.
[a] H. Parameters means Hyper Parameters

B. Performance Comparison with Baseline Methods

Our proposed method E(Q)AGNN-PPIS, have been compared with various baseline methods including GHGPR-PPIS [8], which is the state-of-the-art method for the dataset used. In order to make the comparison fare, we conducted a comparative analysis against a spectrum of established baseline methods including machine learning based methods such as PSIVER [9], ProNA2020 [43] which primarily leverages sequential features, and deep learning based methods like DeepPPISP [16], SPPIDER [44], MaSIF-site [45], which use local and global features of protein sequence either by extracting protein molecular surface fingerprints [45] or directly applying CNNs at the residue level [16]. Furthermore, our evaluation included comparisons with established GNN-based methods such as GraphPPIS [6], AGAT-PPIS [7] and GHGPR-PPIS [8]. These methods construct protein graphs from it's tertiary structures and employ various GNN methods to learn node features: GraphPPIS used GCN [24] with residual connections to capture the local structural information, AGAT-PPIS used GAT [26] with edge features to learn the importance of neighboring nodes and GHGPR-PPIS utilizes GraphHeat [46] with self edge attention mechanism to model the heat diffusion process on the protein graphs. All these methods aim to accurately predict PPI binding sites by exploiting the graph representation of proteins and learning informative node features.

Experimental results presented in Table IV demonstrate the superior performance of our proposed method, E(Q)AGNN-PPIS, in predicting PPI binding sites compared to existing state-of-the-art techniques. E(Q)AGNN-PPIS achieves the highest scores across all evaluation metrics on the test_60 dataset, with an significant improvement of **10%** in MCC (0.55), **6.90%** in F1 score (0.62), **9.68%** in Recall (0.68), **5.45%** in Precision (0.58) and **1.28%** in Accuracy (0.87) when compared to the second best method GHGPR-PPIS. Given the imbalanced nature of the dataset, AUPRC becomes the most crucial metric for evaluation [47]. In this regard, E(Q)AGNN-PPIS demonstrates a significant improvement, achieving a **8.33%** higher AUPRC (0.65) compared to GHGPR-PPIS. The superior performance of E(Q)AGNN-PPIS stems from its implementation of equivariant geometric deep learning with attention mechanism. By incorporating GVP-Module, our architecture preserves crucial 3-D structural information of proteins while processing node and edge features. The attention-based message passing scheme then allows the model to dynamically focus on the most relevant spatial relationships between residues. This combination enables E(Q)AGNN-PPIS to capture both local geometric patterns and long-range interactions critical for accurate PPI site prediction.

C. Generalization and Real-World Deployment Opportunities for E(Q)AGNN-PPIS in Protein Interaction Studies

1) Generalization: To assess the generalization capability of our proposed method, we conducted evaluations on two independent and diverse test sets, namely Test_315-28 and Ubtest_31-6 [7]. Table V presents a comprehensive comparison of our results with current state-of-the-art (SOTA) methods, including GraphPPIS [6], AGAT-PPIS [7], GHGPR-PPIS [8], and several others. Considering the imbalanced nature of the datasets, AUPRC serves as the crucial evaluation metric [47]. The empirical findings demonstrate the robust generalization ability of our method, as it consistently outperforms both AGAT-PPIS and GHGPR-PPIS on the Ubtest_31-6 dataset across all metrics, notably achieving an **15.34%** improvement in AUPRC over AGAT-PPIS and a **14.71%** improvement over GHGPR-PPIS. Moreover, on the Test_315-28 dataset, our method achieves the highest AUPRC, representing a **5.1%** improvement over AGAT-PPIS and a **6.93%** improvement over GHGPR-PPIS, alongside notable enhancements in both F1 and MCC scores. Furthermore, our method demonstrates superior performance compared to earlier methods such as DeepPPISP, SPPIDER, and MaSIF-site, showcasing significant advancements in the field. These outcomes highlight the effectiveness of our approach in handling unseen data inputs and its potential for real-world applications in protein-protein interaction site prediction.

TABLE IV
COMPARATIVE ANALYSIS OF E(Q)AGNN-PPIS AND OTHER BASELINE METHODS ON THE TEST_60 DATASET.

Performance Metric	Machine Learning Based		Deep Learning Based			Graph Machine Learning Based			
	PSIVER [9]	ProNA2020 [43]	DeepPPISP [16]	SPPIDER [44]	MaSIF-site [45]	GraphPPIS [6]	AGAT-PPIS [7]	GHGPR-PPIS [8]	E(Q)AGNN-PPIS
Accuracy	0.56	0.74	0.66	0.75	0.78	0.78	0.86	0.86	**0.87**
Precision	0.19	0.28	0.24	0.33	0.37	0.37	0.54	0.55	**0.58**
Sensitivity (Recall)	0.53	0.40	0.54	0.56	0.56	0.58	0.60	0.62	**0.68**
F1-Score	0.28	0.33	0.34	0.42	0.45	0.45	0.57	0.58	**0.62**
MCC	0.07	0.18	0.17	0.29	0.33	0.33	0.48	0.50	**0.55**
AUPRC	0.19	N/A	0.28	0.37	0.44	0.43	0.57	0.60	**0.65**

The best results are in **bold**.
Note: Prediction results of baseline methods are taken directly from the GHGPR-PPIS [8], and for the ProNA2020 information about AUPRC was not available. For the GraphPPIS, AGAT-PPIS and GHGPR-PPIS methods, we verified the results by reproducing from their provided source code.

TABLE V
GENERALIZATION PERFORMANCE COMPARISON OF E(Q)AGNN-PPIS WITH CURRENT SOTA METHODS.

Method	Test_315-28			Ubtest_31-6		
	F-1	MCC	AUPRC	F-1	MCC	AUPRC
DeepPPISP	N/A	0.169	0.256	N/A	0.162	0.217
SPPIDER	N/A	0.294	0.376	N/A	0.222	0.260
MaSIF-site	N/A	0.304	0.372	N/A	0.141	0.225
GraphPPIS	0.420	0.336	0.423	0.386	0.280	0.323
AGAT-PPIS	0.559	0.481	0.572	0.414	0.327	0.365
GHGPR-PPIS	0.564	0.486	0.562	0.441	0.356	0.367
E(Q)AGNN-PPIS	**0.573**	**0.497**	**0.604**	**0.444**	**0.362**	**0.421**

The best results are in **bold**.
Note: The generalization outcomes for the methods GraphPPIS, AGAT-PPIS, and GHGPR-PPIS have been reproduced using their respective source codes. The results for DeepPPISP, SPPIDER, and MaSIF-site are referenced from GraphPPIS [6], with the F-1 score data not being available.

TABLE VI
COMPARISON OF PREDICTED RESULTS FROM E(Q)AGNN-PPIS ON PROTEIN '1WPX(A)' OF LENGTH 421 WITH STATE-OF-THE-ART METHODS.

Method	Predicted Binding Sites	TP	FP	FN	F1-Score
AGAT-PPIS	28	18	10	28	49%
GHGPR-PPIS	60	33	27	13	62%
E(Q)AGNN-PPIS	33	32	1	14	**81%**

The best results are in **bold**.

2) Real-World Deployment Opportunities for E(Q)AGNN-PPIS: To assess the practical utility of the E(Q)AGNN-PPIS model, we undertook a case study analyzing individual proteins selected from the Test_60 dataset. Specifically, we examined 10 proteins: 3cf4(A), 4cdg(A), 3zhe(B), 1wpx(A), 3fmo(A), 2fhz(B), 3mcb(A), 3vz9(D), 4nqw(B) and 4je3(B), with the letters in parentheses indicating the specific protein chains. We have selected these proteins carefully by choosing 5 longest and 5 shortest proteins from the dataset to represent diversity. We present our observations particularly on the 1wpx(A) protein chain, which is one of the longest in the dataset, consisting of 421 amino acid residues including 46 labelled as binding sites. Detailed observations for the other proteins are documented in Appendix A.

For the protein chain 1wpx(A), we compared the E(Q)AGNN-PPIS with two SOTA models, GHGPR-PPIS and AGAT-PPIS, focusing on their ability to accurately identify binding sites. As per the results summarized in Table VI, the E(Q)AGNN-PPIS identified 33 residues as binding sites, of which 32 were confirmed as true positives and only one as a false positive, with 14 residues as false negative. In contrast, AGAT-PPIS and GHGPR-PPIS predicted significantly less and more binding sites, respectively and both the methods achieved F1-Score of **49%** and **62%**. Notably, our method achieved a superior F1-Score of **81%**, demonstrating its effectiveness in identifying potential binding sites within complex protein chains. The results exhibits the enhanced capability of E(Q)AGNN-PPIS to not only predict with high precision but also to maintain a lower rate of false positives, thus providing a robust tool for the prediction of PPI binding sites. In Figure 3, we present the comparative visualizations of the predictions from the GHGPR-PPIS 3c and E(Q)AGNN-PPIS 3b methods for the protein '1wpx(A)'. The results indicate that the predictions made using E(Q)AGNN-PPIS align more closely with the actual observations, showcasing its enhanced accuracy in PPI binding sites prediction.

3) Length-invariant Performance of E(Q)AGNN-PPIS: To evaluate the effectiveness of E(Q)AGNN-PPIS across varying protein lengths, we categorized the Test_60 dataset into small (52-127 residues), medium (128-247 residues), and large (253-766 residues) proteins. As shown in Table VII, E(Q)AGNN-PPIS consistently outperforms AGAT-PPIS and GHGPR-PPIS across all length categories, achieving average F1 scores of 0.672, 0.634, and 0.563 for small, medium, and large proteins, respectively. This robust performance, with improvements of up to **32.15%** over AGAT-PPIS and **10.3%** over GHGPR-PPIS, demonstrates E(Q)AGNN-PPIS's length-invariant predictive capability. The model's ability to maintain high accuracy across

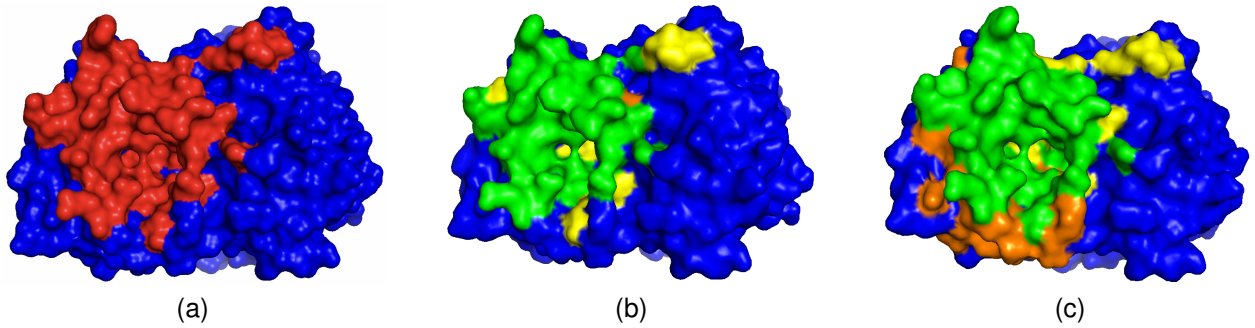

Fig. 3. An example case of protein '1wpx(A)', (a) original binding site, (b) prediction using our method E(Q)AGNN-PPIS, and (c) prediction using GHGPR-PPIS. We used color codes Red, Green, Orange, Yellow and Blue to indicate original binding sites, True Positive, False Positive, False Negative and non binding sites, respectively. Visualization is done using PyMOL [48] tool.

diverse protein sizes can be attributed to its effective utilization of geometric equivariance properties and attention mechanisms, making it well-suited for real-world binding site prediction tasks.

TABLE VII
AVERAGE F1 SCORES FOR DIFFERENT PROTEIN LENGTHS ACROSS MODELS

Protein Length	E(Q)AGNN-PPIS	AGAT-PPIS	GHGPR-PPIS
Small	**0.672**	0.599	0.639
Medium	**0.634**	0.599	0.599
Large	**0.563**	0.426	0.510

D. Ablation Studies

To demonstrate the contribution of each component of our pipeline to the overall performance of PPI binding site prediction, we have performed an ablation study through additional experiments.

1) Evaluating the Effect of Threshold (θ) on protein distance map: As mentioned in section II-B1 we derived the adjacency matrix by using a threshold θ on the protein distance matrix; as shown in Equation 1. Here we report the performance of E(Q)AGNN-PPIS at various distance thresholds on Test_60 dataset.

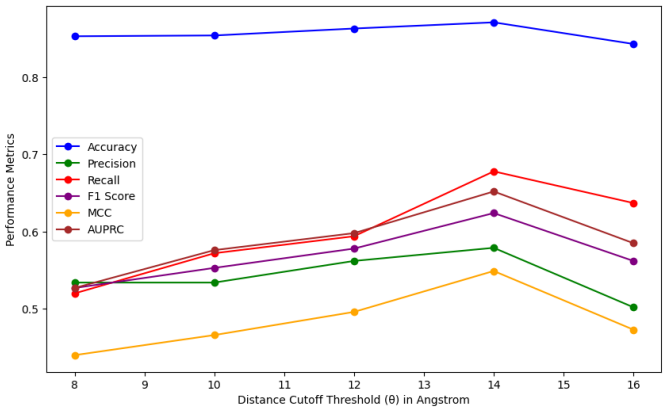

Fig. 4. Performance of E(Q)AGNN-PPIS on different distance cutoff threshold θ. Various matrix are presented in different colours.

From the Figure 4, it can be seen that as the distance cutoff threshold increases, the model's performance generally improves across all metrics because information flow increases, until reaching a peak at around 14Å. Beyond this point, the performance metrics experience a slight decline. As the cutoff increases, the protein graph becomes increasingly dense, leading to oversquashing [49], which occurs when long-range dependencies in dense graphs cannot be effectively captured, resulting in information loss. The optimal 14Å cutoff thus represents a balance between sufficient information propagation and preservation of meaningful geometric relationships in the protein structure. The Accuracy and AUPRC exhibit the highest values among the metrics, followed by F1 Score, MCC, Precision, and Recall. Throughout this work we have utilized the distance cutoff of 14Å.

2) Effect of number of Layers:: This section investigates the influence of model depth on the performance of the E(Q)AGNN-PPIS model. Figure 5 illustrates the effect of varying the number of E(Q)AGNN layers on a range of evaluation metrics, using the Test_60 dataset. The results demonstrate that increasing

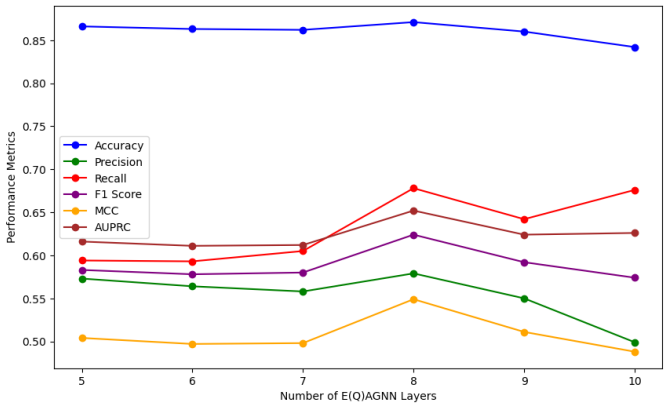

Fig. 5. Performance of E(Q)AGNN-PPIS on different number of E(Q)AGNN layers.

the number of layers from 5 to 10 generally enhances the model's performance, with Accuracy and AUPRC exhibiting the highest values and the most consistent improvement. However, the optimal number of layers is dependent on the specific metric and the trade-off between performance gain

and computational complexity. In this study, we employed 8 layers across all experiments, as it provides a favorable balance between performance and computational efficiency.

V. Conclusion

In this research, we introduce E(Q)AGNN-PPIS, an attention-enhanced equivariant graph neural network specifically designed for predicting protein binding sites without prior knowledge of binding partners. Our methodology integrates an innovative attention mechanism that processes both scalar and vector features of the protein graph, facilitating effective node representation learning. By incorporating eight layers of E(Q)AGNN with residual connections, our model demonstrates superior robustness and generalization capabilities compared to other state-of-the-art methods across diverse test datasets. While our approach incorporates structural features, evolutionary information, secondary structure level information, and atomic level features, we acknowledge that integrating additional physico-chemical properties (such as dissociation constants and optical activity etc.) could potentially enhance prediction accuracy. Moreover, our current method focuses on unilateral binding site prediction. A natural extension of this work would be to leverage binding partner information to identify complementary binding sites, considering both geometric and physico-chemical compatibility. This approach could pave the way for modeling protein dimeric complex structures, significantly advancing our understanding of protein-protein interactions.

Looking ahead, we plan to expand this work by developing a model that more effectively addresses these limitations while maintaining adherence to the geometric intricacies of protein data. We are optimistic that our approach could be readily adapted to a broader range of applications, including the prediction of protein interactions with small ligands, DNA, and RNA, thereby contributing to the development of a more generalized model for biomolecular interaction prediction.

References

[1] E. D. Scheeff and J. L. Fink, "Fundamentals of protein structure," *Structural bioinformatics*, vol. 44, pp. 15–39, 2003.

[2] N. J. Burgoyne and R. M. Jackson, "Predicting protein interaction sites: Binding hot-spots in protein-protein and protein-ligand interfaces," *Bioinformatics*, vol. 22, no. 11, pp. 1335–1342, 2006.

[3] R. B. Russell and P. Aloy, "Targeting and tinkering with interaction networks," *Nature chemical biology*, vol. 4, no. 11, pp. 666–673, 2008.

[4] S. Ghosh and P. Mitra, "Matpip: A deep-learning architecture with explainable ai for sequence-driven, feature mixed protein-protein interaction prediction," *Computer Methods and Programs in Biomedicine*, vol. 244, p. 107 955, 2024.

[5] X. Wang, B. Yu, A. Ma, C. Chen, B. Liu, and Q. Ma, "Protein–protein interaction sites prediction by ensemble random forests with synthetic minority oversampling technique," *Bioinformatics*, vol. 35, no. 14, pp. 2395–2402, 2019.

[6] Q. Yuan, J. Chen, H. Zhao, Y. Zhou, and Y. Yang, "Structure-aware protein–protein interaction site prediction using deep graph convolutional network," *Bioinformatics*, vol. 38, no. 1, pp. 125–132, 2022.

[7] Y. Zhou, Y. Jiang, and Y. Yang, "AGAT-PPIS: A novel protein–protein interaction site predictor based on augmented graph attention network with initial residual and identity mapping," *Briefings in Bioinformatics*, vol. 24, no. 3, bbad122, 2023.

[8] X. Zeng, F.-F. Meng, X. Li, K.-Y. Zhong, B. Jiang, and Y. Li, "GHGPR-PPIS: A graph convolutional network for identifying protein-protein interaction site using heat kernel with generalized pagerank techniques and edge self-attention feature processing block," *Computers in Biology and Medicine*, vol. 168, p. 107 683, 2024.

[9] Y. Murakami and K. Mizuguchi, "Applying the naïve bayes classifier with kernel density estimation to the prediction of protein–protein interaction sites," *Bioinformatics*, vol. 26, no. 15, pp. 1841–1848, 2010.

[10] N. Li, Z. Sun, and F. Jiang, "Prediction of protein-protein binding site by using core interface residue and support vector machine," *BMC bioinformatics*, vol. 9, pp. 1–13, 2008.

[11] P. Fariselli, F. Pazos, A. Valencia, and R. Casadio, "Prediction of protein–protein interaction sites in heterocomplexes with neural networks," *European Journal of Biochemistry*, vol. 269, no. 5, pp. 1356–1361, 2002.

[12] J. Zhang and L. Kurgan, "SCRIBER: Accurate and partner type-specific prediction of protein-binding residues from proteins sequences," *Bioinformatics*, vol. 35, no. 14, pp. i343–i353, 2019.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[14] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.

[15] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" *arXiv preprint arXiv:1810.00826*, 2018.

[16] M. Zeng, F. Zhang, F.-X. Wu, Y. Li, J. Wang, and M. Li, "Protein–protein interaction site prediction through combining local and global features with deep neural networks," *Bioinformatics*, vol. 36, no. 4, pp. 1114–1120, 2020.

[17] J. Jiménez, S. Doerr, G. Martínez-Rosell, A. S. Rose, and G. De Fabritiis, "DeepSite: Protein-binding site predictor using 3d-convolutional neural networks," *Bioinformatics*, vol. 33, no. 19, pp. 3036–3042, 2017.

[18] Y. Li, G. B. Golding, and L. Ilie, "DELPHI: Accurate deep ensemble model for protein interaction sites prediction," *Bioinformatics*, vol. 37, no. 7, pp. 896–904, 2021.

[19] Y. Kang, Y. Xu, X. Wang, *et al.*, "HN-PPISP: A hybrid network based on mlp-mixer for protein–protein interaction site prediction," *Briefings in Bioinformatics*, vol. 24, no. 1, bbac480, 2023.

[20] J. Kandel, H. Tayara, and K. T. Chong, "PUResNet: Prediction of protein-ligand binding sites using deep residual neural network," *Journal of cheminformatics*, vol. 13, pp. 1–14, 2021.

[21] B. Zhang, J. Li, L. Quan, Y. Chen, and Q. Lü, "Sequence-based prediction of protein-protein interaction sites by simplified long short-term memory network," *Neurocomputing*, vol. 357, pp. 86–100, 2019.

[22] Y. Zhang, W. Huang, Z. Wei, Y. Yuan, and Z. Ding, "Equipocket: An e (3)-equivariant geometric graph neural network for ligand binding site prediction," *arXiv preprint arXiv:2302.12177*, 2023.

[23] J. Jumper, R. Evans, A. Pritzel, *et al.*, "Highly accurate protein structure prediction with alphafold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[24] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proceedings of the 5th International Conference on Learning Representations, ICLR*, 2017.

[25] S. Wang, W. Chen, P. Han, X. Li, and T. Song, "RGN: Residue-based graph attention and convolutional network for protein–protein interaction site prediction," *Journal of chemical information and modeling*, vol. 62, no. 23, pp. 5961–5974, 2022.

[26] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proceedings of the 6th International Conference on Learning Representations, ICLR*, 2018.

[27] V. G. Satorras, E. Hoogeboom, and M. Welling, "E (n) equivariant graph neural networks," in *International conference on machine learning*, PMLR, 2021, pp. 9323–9332.

[28] B. Jing, S. Eismann, P. Suriana, R. J. L. Townshend, and R. Dror, "Learning from protein structure with geometric vector perceptrons," in *International Conference on Learning Representations*, 2020.

[29] J. Bernett, D. B. Blumenthal, and M. List, "Cracking the black box of deep sequence-based protein–protein interaction prediction," *Briefings in Bioinformatics*, vol. 25, no. 2, bbae076, 2024.

[30] F. Madeira, N. Madhusoodanan, J. Lee, *et al.*, "The embl-ebi job dispatcher sequence analysis tools framework in 2024," *Nucleic Acids Research*, gkae241, 2024.

[31] B. E. Suzek, H. Huang, P. McGarvey, R. Mazumder, and C. H. Wu, "Uniref: Comprehensive and non-redundant uniprot reference clusters," *Bioinformatics*, vol. 23, no. 10, pp. 1282–1288, 2007.

[32] M. Remmert, A. Biegert, A. Hauser, and J. Söding, "HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment," *Nature methods*, vol. 9, no. 2, pp. 173–175, 2012.

[33] W. Kabsch and C. Sander, "Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features," *Biopolymers: Original Research on Biomolecules*, vol. 22, no. 12, pp. 2577–2637, 1983.

[34] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *International conference on machine learning*, PMLR, 2017, pp. 1263–1272.

[35] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," *arXiv preprint arXiv:2104.13478*, 2021.

[36] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE transactions on neural networks*, vol. 20, no. 1, pp. 61–80, 2008.

[37] B. Jing, S. Eismann, P. N. Soni, and R. O. Dror, "Equivariant graph neural networks for 3d macromolecular structure," *arXiv preprint arXiv:2106.03843*, 2021.

[38] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.

[39] A. Paszke, S. Gross, F. Massa, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.

[40] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," *arXiv preprint arXiv:1903.02428*, 2019.

[41] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[42] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, "Simple and deep graph convolutional networks," in *International conference on machine learning*, PMLR, 2020, pp. 1725–1735.

[43] J. Qiu, M. Bernhofer, M. Heinzinger, *et al.*, "ProNA2020 predicts protein–dna, protein–rna, and protein–protein binding proteins and residues from sequence," *Journal of molecular biology*, vol. 432, no. 7, pp. 2428–2443, 2020.

[44] A. Porollo and J. Meller, "Prediction-based fingerprints of protein–protein interactions," *Proteins: Structure, Function, and Bioinformatics*, vol. 66, no. 3, pp. 630–645, 2007.

[45] P. Gainza, F. Sverrisson, F. Monti, *et al.*, "Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning," *Nature Methods*, vol. 17, no. 2, pp. 184–192, 2020.

[46] B. Xu, H. Shen, Q. Cao, K. Cen, and X. Cheng, "Graph convolutional networks using heat kernel for semi-supervised learning," *arXiv preprint arXiv:2007.16002*, 2020.

[47] J. T. Hancock, T. M. Khoshgoftaar, and J. M. Johnson, "Evaluating classifier performance with highly imbalanced big data," *Journal of Big Data*, vol. 10, no. 1, p. 42, 2023.

[48] W. L. DeLano *et al.*, "Pymol: An open-source molecular graphics tool," *CCP4 Newsl. Protein Crystallogr*, vol. 40, no. 1, pp. 82–92, 2002.

[49] M. Black, Z. Wan, A. Nayyeri, and Y. Wang, "Understanding oversquashing in gnns through the lens of effective resistance," in *International Conference on Machine Learning*, PMLR, 2023, pp. 2528–2547.

APPENDIX A
ABLATION RESULTS

Here we provide the full prediction details of selected protein from our method, E(Q)AGNN-PPIS as explained in Section IV-C2 and provide the comparison with the state-of-the-art method GHGPR-PPIS.

TABLE VIII
PREDICTED RESULTS FROM E(Q)AGNN-PPIS AND GHGPR-PPIS METHOD ON VARIOUS PROTEINS OF DIFFERENT LENGTHS.

Protein-ID[a]	Length	Method	Predicted binding sites	TP	FP	FN	F1-Score
3cf4(A)	766	Ours	112	36	76	21	43%
		GHGPR-PPIS	98	35	63	22	**45%**
4cdg(A)	625	Ours	69	21	48	18	**39%**
		GHGPR-PPIS	99	18	81	21	26%
3fmo(A)	421	Ours	34	23	11	11	**68%**
		GHGPR-PPIS	25	17	8	17	58%
3zhe(B)	390	Ours	42	26	16	11	**66%**
		GHGPR-PPIS	55	26	29	11	57%
3mcb(A)	54	Ours	40	23	17	6	67%
		GHGPR-PPIS	48	27	21	2	**70%**
3vz9(D)	60	Ours	31	25	6	8	**78%**
		GHGPR-PPIS	33	25	8	8	76%
4nqw(B)	77	Ours	39	38	1	6	**92%**
		GHGPR-PPIS	36	32	4	12	80%
4je3(B)	73	Ours	34	30	4	7	**85%**
		GHGPR-PPIS	44	29	15	8	72%
2fhz(B)	93	Ours	43	31	12	1	**83%**
		GHGPR-PPIS	29	16	13	16	52%

[a] In Protein-ID column, it follows the structure PDB-ID(chain).
'Ours' refers to E(Q)AGNN-PPIS method.
Best method values are in **bold**.

APPENDIX B
PERFORMANCE METRICS

$$\text{Accuracy (ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{13}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

$$\text{Recall} = \textit{Sensitivity} = \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{15}$$

$$\text{F1} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{16}$$

$$\text{MCC} = \frac{\text{TP} * \text{TN} - \text{FP} * \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{17}$$

In the above metrics, TP (True Positive) denotes the number of correctly identified PPI binding sites, whereas TN (True Negative) refers to the correct predictions of non-PPI binding sites. Conversely, FP (False Positive) indicates instances where non-binding sites are erroneously predicted as binding sites, and FN (False Negative) marks actual binding sites that are misclassified as non-binding. Additionally, TPR (True Positive Rate), also known as sensitivity, quantifies the ratio of actual PPI binding sites that are accurately detected by the model.

APPENDIX C
PROOFS

A. Proof of Proposition 1

Proposition 1: Let $\vec{x_i} \in \mathbb{R}^3$ be the position vector of the ith node representing an amino acid, and $\mathcal{N}_i$ be the set of neighbors of node i. The node vector feature $\vec{h_v}^i$ defined as:

$$\vec{h_v}^i = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} (\vec{x_j} - \vec{x_i}) \tag{18}$$

is invariant under translations and equivariant under the orthogonal group O(3).

Proof. We prove the invariance under translations and equivariance under O(3) separately:

Case 1. Translation: Under a translation $\vec{b}$, the positions transform as $\vec{x_j} \to \vec{x_j} + \vec{b}$ and $\vec{x_i} \to \vec{x_i} + \vec{b}$. The transformed $\vec{h_v}^i$ is:

$$\begin{aligned} \vec{h'_v}^i &= \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} ((\vec{x_j} + \vec{b}) - (\vec{x_i} + \vec{b})) \\ &= \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} (\vec{x_j} - \vec{x_i}) = \vec{h_v}^i \end{aligned}$$

This demonstrates translation invariance of $\vec{h_v}^i$.

Case 2. O(3) Transformation: For an orthogonal matrix $R \in \mathbb{R}^{3\times3}$, vectors transform as $\vec{x_j} \to R\vec{x_j}$ and $\vec{x_i} \to R\vec{x_i}$. The transformed $\vec{h_v}^i$ is:

$$\begin{aligned} \vec{h'_v}^i &= \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} (R\vec{x_j} - R\vec{x_i}) \\ &= \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} R(\vec{x_j} - \vec{x_i}) \\ &= R\left(\frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} (\vec{x_j} - \vec{x_i})\right) = R\vec{h_v}^i \end{aligned}$$

This demonstrates equivariance of $\vec{h_v}^i$ under O(3) transformations. □

B. Proof of Proposition 2

Proposition 2: Let $\vec{x_i}, \vec{x_j} \in \mathbb{R}^3$ be the position vectors of nodes i and j. The edge vector feature $\vec{h_e}^{j\to i}$ defined as:

$$\vec{h_e}^{j\to i} = \frac{\vec{x_j} - \vec{x_i}}{\|\vec{x_j} - \vec{x_i}\|} \tag{19}$$

is invariant under translations and equivariant under orthogonal group O(3).

Proof. We prove the invariance under translations and equivariance under O(3) separately:

Case 1. Translation: Under a translation $\vec{b}$, the positions transform as $\vec{x_j} \to \vec{x_j} + \vec{b}$ and $\vec{x_i} \to \vec{x_i} + \vec{b}$. The transformed $\vec{h_e}^{j\to i}$ is:

$$\begin{aligned}\vec{h'_e}^{j\to i} &= \frac{(\vec{x_j} + \vec{b}) - (\vec{x_i} + \vec{b})}{\|(\vec{x_j} + \vec{b}) - (\vec{x_i} + \vec{b})\|} \\ &= \frac{\vec{x_j} - \vec{x_i}}{\|\vec{x_j} - \vec{x_i}\|} = \vec{h_e}^{j\to i}\end{aligned}$$

This demonstrates translation invariance of $\vec{h_e}^{j\to i}$.

Case 2. O(3) Transformation: For an orthogonal matrix $R \in \mathbb{R}^{3\times 3}$, vectors transform as $\vec{x_j} \to R\vec{x_j}$ and $\vec{x_i} \to R\vec{x_i}$. The transformed $\vec{h_e}^{j\to i}$ is:

$$\begin{aligned}\vec{h'_e}^{j\to i} &= \frac{R\vec{x_j} - R\vec{x_i}}{\|R\vec{x_j} - R\vec{x_i}\|} \\ &= \frac{R(\vec{x_j} - \vec{x_i})}{\|R(\vec{x_j} - \vec{x_i})\|} \\ &= \frac{R(\vec{x_j} - \vec{x_i})}{\|\vec{x_j} - \vec{x_i}\|} \quad \text{(since } R \text{ preserves norms)} \\ &= R\left(\frac{\vec{x_j} - \vec{x_i}}{\|\vec{x_j} - \vec{x_i}\|}\right) = R\vec{h_e}^{j\to i}\end{aligned}$$

This demonstrates equivariance of $\vec{h_e}^{j\to i}$ under O(3) transformations.

□

C. Proof of Proposition 3

Proposition 3: We proof that the E(Q)AGNN-PPIS model, including its attention mechanism as defined in Equations (5)–(9), is equivariant under the actions of the orthogonal group $O(3)$.

Proof. Let $R \in \mathbb{R}^{3\times 3}$ be an orthogonal metrix.

Base case: Input features The input vector features $\vec{h_v}^{i}$ for any node i and $\vec{h_e}^{j\to i}$ for the edge between nodes j and i are equivariant under any transformation $R \in O(3)$ as established in Proposition 1, 2 and scalar features are invariant by definition.

Inductive step: These input features are used in subsequent layers of our method, we prove that operations in layer $t+1$ preserve equivariance as input features at layer t are equivariant.

1) GVP-Module: The GVP-Module preserves equivariance as shown in [37]. For completeness:

$$\text{GVP}(s, R(\vec{v})) = (s', R(\vec{v}')) \text{ when } \text{GVP}(s, \vec{v}) = (s', \vec{v}')$$

2) Attention Mechanism: The attention scores α_{ij} are computed using only scalar features:

$$\beta_{ij} = \text{LeakyReLU}(a^T(W^\alpha[W^v h_v^i || W^v h_v^j]))$$

$$\alpha_{ij} = \text{softmax}(\beta_{ij} \cdot A_{ij})$$

Since scalar features and the adjacency matrix A_{ij} are invariant under any transformation R , α_{ij} is also invariant.

3) Message Passing: The message function is:

$$m_{ij} = \mathcal{S}(\mathfrak{h}_\mathfrak{v}^i || \mathfrak{h}_\mathfrak{e}^{j\to i})$$

where $\mathcal{S}$ is a sequence of GVP-Modules. Since GVP-Modules preserve equivariance, and concatenation of equivariant features remains equivariant, m_{ij} is equivariant.

4) Feature Update: The feature update is:

$$\mathfrak{h}_\mathfrak{v}^i \leftarrow \text{LayerNorm}(\mathfrak{h}_\mathfrak{v}^i + \frac{1}{k_i}\text{Dropout}(\sum_{j:j\neq i} \alpha_{ij} m_{ij}))$$

Since α_{ij} is invariant (scalar), m_{ij} is equivariant and sum of equivariant features is equivariant, therefore, the updated features $\mathfrak{h}_\mathfrak{v}^i$ are equivariant.

By induction, all layers of the network preserve equivariance. Therefore, the entire E(Q)AGNN-PPIS model is equivariant under $O(3)$ transformations.

□

Animesh received the B.Tech degree in Electrical and Electronics Engineering from the Pranveer Singh Institute of Technology, Kanpur, India, in 2016 and M.tech degree in Electrical Engineering with specialization in Control System from the National Institute of Technology (NIT), Kurukshetra, India in 2020. He is currently pursuing a Ph.D. in the Department of Artificial Intelligence at the Indian Institute of Technology (IIT), Kharagpur. His research interests include graph machine learning and its applications.

Rishi Suvvada pursued an interdisciplinary Dual degree with B.Tech in Chemical Engineering and M.Tech in Artificial intelligence and Machine learning from the Department of Artificial Intelligence at Indian Institute of Technology Kharagpur,India from 2019 to 2024. He is currently pursuing his corporate stint working as a Data Scientist in a Heath Tech startup called Hilabs His research interests include graph machine learning and its applications, AI in medicine and finance space.

Plaban Kumar Bhowmick received the B.Tech degree from University of Calcutta, India, 2002, and the master's and Ph.D. degrees in Computer Science and Engineering from the Indian Institute of Technology (IIT) Kharagpur, India, in 2006 and 2011, respectively. He is an Associate Professor with the Department of Artificial Intelligence, IIT Kharagpur. He has been Co-Principal Investigator and Technical Lead of the National Digital Library of India project. His research areas include text processing, artificial intelligence in education, semantic web technology, knowledge graphs, graph machine learning, and digital library technologies.

Pralay Mitra received the BSc degree with honors in physics and BTech degree in Computer Science and Engineering from the University of Calcutta, India, in 1999 and in 2002, respectively. He finished the ME degree in computer science and engineering from the Bengal Engineering and Science University, Shibpur (currently known as IIEST, Shibpur), India, in 2004, and the PhD degree from the Indian Institute of Science, Bangalore, in 2010. He was the research associate with the Indian Institute of Science, Bangalore (from 2010 to 2011) and the research fellow with the University of Michigan, Ann Arbor (from 2011-2013). Since 2013, he has been at the Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India as an Assistant Professor (2013-2020) and then as an Associate Professor (2020-present). He is actively working on bioinformatics and computational biology. Over the years, he has developed expertise in modeling and designing protein structures and protein functions. He has developed a number of algorithms for protein-protein docking, predicting protein assembly from the crystal structures, and protein design.