Quantum Machine Intelligence manuscript No.
(will be inserted by the editor)

Q-CVaR-Fold: A Quantum CVaR Contrastive Learning Framework for Protein-Folding Decoy Scoring

Parham Ghayour

Received: date / Accepted: date

Abstract Protein structure prediction and decoy ranking remain central challenges in computational biophysics. Classical scoring functions often struggle to discriminate near-native conformations from large populations of plausible decoys, particularly in the critical low-energy tail of the conformational distribution. We introduce Q-CVaR-Fold, a hybrid quantum–classical architecture that integrates a geometric graph neural network encoder with a small parameterized quantum circuit acting as a nonlinear scoring head. To focus optimization on near-native conformations, we combine contrastive ranking with Conditional Value-at-Risk (CVaR) tail reweighting, yielding a risk-sensitive training objective aligned with structural evaluation metrics.

Despite using only four qubits and shallow entangling layers, Q-CVaR-Fold exhibits stable end-to-end training and avoids barren plateaus. On a decoy-ranking benchmark, the model achieves a ROC-AUC of 0.984 and perfect top-5 native enrichment across all sequences, outperforming classical baselines of comparable size. The score distributions and monotonic reduction of CVaR loss demonstrate that quantum feature transformations, coupled with tail-focused optimization, provide discriminative power beyond standard MLP heads. To our knowledge, this is the first demonstration of a quantum-enhanced, risk-sensitive scoring model that achieves near-perfect recovery of native structures in decoy-ranking tasks.

Q-CVaR-Fold highlights the potential of hybrid quantum models for energy landscape modeling, fragment selection, and structural refinement, and offers a promising foundation for next-generation quantum–geometric methods in computational structural biology.

Keywords Quantum machine learning · Protein folding · Decoy scoring · Conditional Value-at-Risk (CVaR) · Contrastive learning · Parameterized quantum circuits

Parham Ghayour
Independent Researcher,
E-mail: monsieur.parham.ghayour@gmail.com

1 Introduction

Predicting the three-dimensional structure of a protein and evaluating the quality of candidate conformations remain central challenges in structural biology. Although recent advances in deep learning have dramatically improved sequence-to-structure prediction, particularly with the emergence of transformer-based architectures, downstream decoy scoring continues to play a key role in fragment assembly, structural refinement, and quality assessment pipelines. In these settings, the goal is not merely to identify grossly incorrect structures, but to reliably discriminate near-native conformations from large sets of plausible but incorrect decoys. This task is intrinsically challenging because the relevant decoys often lie in a narrow band of low-energy conformations, where small geometric perturbations can produce significant ambiguity in conventional scoring functions.

Classical approaches to decoy ranking, such as physics-based energy models, fragment-based likelihoods, or geometric graph neural networks, have achieved substantial progress [1–3, 12]. Nevertheless, even state-of-the-art scoring functions may exhibit limited sensitivity in the near-native regime, where the majority of evaluation metrics—including GDT-TS and TM-score [5]—are most discerning. In this context, methods that can amplify subtle geometric differences or reparameterize conformation space into a more separable representation are of particular interest.

Quantum machine learning (QML) has recently emerged as a promising direction for biophysics and molecular modeling [6]. Parameterized quantum circuits (PQCs) can implement expressive nonlinear transformations in high-dimensional Hilbert spaces, potentially offering inductive biases distinct from those available to classical neural networks [8, 15]. Several works have explored quantum kernel methods, VQE-based energy estimators, and quantum graph constructions for molecular learning, though applications to protein folding remain limited [19, 21]. Existing methods typically target energy regression, rather than ranking, and do not explicitly emphasize the rare near-native decoys that dominate structure evaluation tasks.

In parallel, risk-sensitive objectives based on Conditional Value-at-Risk (CVaR) have shown promise in problems where performance on the worst-case or tail distribution is critical [13,14]. Protein decoy scoring exhibits this exact structure: the hardest decoys—those closest to the native fold—are the ones that matter most, both for structure prediction and refinement. Yet CVaR and related tail-aware methods have rarely been applied to molecular structure evaluation or to QML architectures.

In this work, we introduce **Q-CVaR-Fold**, a hybrid quantum–classical model designed specifically for ranking protein conformations. Our approach combines three components:

1. a geometric graph neural network that encodes residue-level features and contact patterns into compact embeddings;
2. a parameterized quantum circuit that serves as a nonlinear scoring head, providing a compact but expressive transformation of conformation embeddings; and
3. a contrastive ranking objective augmented with CVaR tail reweighting, focusing optimization on near-native and structurally challenging decoys.

This combination enables Q-CVaR-Fold to reparameterize conformational geometry via quantum feature transformations while explicitly prioritizing accuracy in the near-native tail of the decoy distribution. Despite using only a shallow four-qubit circuit, our method achieves near-perfect discrimination on a decoy-ranking benchmark, including a ROC-AUC of 0.984 and perfect top-5 native enrichment. These results surpass those of classical models of comparable size and demonstrate how risk-sensitive quantum scoring functions can provide meaningful advantages in structural biology tasks.

To our knowledge, this is the first demonstration of a hybrid quantum-enhanced, CVaR-driven ranking model for protein structure evaluation. The proposed architecture offers a compact alternative to large classical networks and suggests a new direction for integrating quantum circuits with geometric learning in molecular modeling.

2 Background

Protein structure prediction aims to identify the three-dimensional conformation that a given amino-acid sequence adopts under physiological conditions. This problem is fundamentally challenging because the search space of possible conformations grows exponentially with sequence length, and the mapping from sequence to structure is governed by a highly rugged and high-dimensional energy landscape. Classical algorithms often approximate this landscape with coarse-grained statistical potentials, fragment-based sampling, or deep neural network models trained on large structural datasets. While recent breakthroughs such as AlphaFold have dramatically improved the field, structure refinement and fine-grained ranking of candidate conformations—known as the *decoy scoring problem*—remain non-trivial, especially for small or poorly constrained protein fragments.

In this section, we review (i) the role of protein decoy scoring in computational biology, (ii) current quantum machine learning approaches for molecular systems and their limitations, and (iii) the motivation for risk-sensitive learning using the Conditional Value-at-Risk (CVaR) objective. These three components provide the conceptual foundation for our proposed Q-CVaR-Fold framework.

2.1 Protein Folding and Decoy Scoring

Given a sequence and a set of candidate conformations (decoys), the goal of decoy scoring is to identify which structures are closest to the native fold. This is a crucial component of many classical folding pipelines, since sampling algorithms—whether fragment assembly, physics-based molecular dynamics, or generative models—tend to produce large collections of plausible but mostly incorrect 3D conformations. A scoring function must therefore assign higher values to near-native conformations and lower values to high-energy, physically implausible structures.

Traditionally, decoy scoring functions rely on combinations of knowledge-based energy potentials, residue-contact statistics, physico-chemical terms, or geometric constraints. Modern approaches often use machine learning models, including graph neural networks (GNNs) that operate on residue-level contact graphs. However, these models typically optimize average-case metrics such as mean squared

error (MSE) on predicted energies or RMSD. Such metrics do not explicitly emphasize the rare, low-energy, near-native conformations that are most relevant biologically. This creates a mismatch between the optimization target and the actual goal of decoy ranking, which is inherently a *tail-sensitive* problem.

2.2 Quantum Machine Learning for Molecular Systems

Quantum machine learning (QML) has emerged as a promising paradigm for modeling molecular systems, largely due to the natural correspondence between quantum states and many-body fermionic systems. Parameterized quantum circuits (PQCs) have been used in variational quantum eigensolvers for estimating ground-state energies, in quantum-enhanced kernel methods, and in quantum versions of graph-convolutional architectures. Several works have also applied quantum optimization or quantum annealing to simplified protein folding models, such as lattice-based HP models, where the folding problem can be formulated as an Ising-type Hamiltonian.

Despite these advances, most QML approaches for biophysical systems remain focused on *energy regression*, i.e., fitting a quantum circuit to reproduce the ground-state energy of an approximate Hamiltonian. While such methods are valuable for studying small molecules, they do not directly address the decoy-ranking task, nor do they incorporate information about the structure of the folding landscape. Furthermore, energy-based models often require accurate Hamiltonian definitions, which are difficult to construct for protein-scale systems without strong approximations.

By contrast, learning a scoring function directly from data using a PQC—without relying on a physical Hamiltonian—provides a more flexible approach that aligns more closely with protein structural objectives. This creates an opportunity for hybrid quantum–classical architectures that combine graph-based representations with quantum non-linearities.

2.3 Risk-Sensitive Learning and the CVaR Objective

The Conditional Value-at-Risk (CVaR) is a well-established risk measure originating from financial mathematics. For a given loss distribution, CVaR at level α corresponds to the average loss in the worst α-fraction of outcomes. Formally, letting L denote a random loss variable with cumulative distribution function F, the CVaR is

$$\mathrm{CVaR}_\alpha(L) = \mathbb{E}[\, L \,|\, L \geq \mathrm{VaR}_\alpha(L)] \,, \tag{1}$$

where VaR_α is the Value-at-Risk threshold.

In machine learning, CVaR provides a principled way to emphasize rare, high-impact errors rather than average-case performance. This is particularly relevant when the objectives involve detecting difficult or informative examples that reside in the tail of the loss distribution. Recent advances in quantum optimization have successfully applied CVaR-based objectives to variational quantum algorithms such as VQE and QAOA, demonstrating improved convergence for rugged or multimodal energy landscapes.

Protein decoy scoring is inherently a tail-sensitive task: the near-native conformations occupy a tiny fraction of conformational space but dominate biological relevance. A risk-neutral objective such as MSE does not sufficiently highlight these critical states. By contrast, a CVaR-based loss naturally prioritizes ranking the near-native structures correctly, enabling the model to learn a scoring function that captures the sharp variations between low-energy basins and the bulk of high-energy decoys.

These considerations motivate the development of a quantum risk-sensitive decoy scoring model, where a PQC is trained using a CVaR-weighted contrastive ranking objective. This forms the conceptual basis of the Q-CVaR-Fold architecture introduced in the next section.

3 Method: Q-CVaR-Fold

3.1 Protein Graph Representation

To apply a graph-based quantum–classical architecture to protein conformations, each candidate structure (native or decoy) must be transformed into a residue-level graph encoding both sequential and spatial information. Given a protein conformation X with residues $\{1, \ldots, N\}$ and Cartesian coordinates $\{\mathbf{r}_i\}$, we construct an undirected graph $G = (V, E)$ whose nodes represent amino acids and whose edges capture covalent connectivity and three-dimensional proximity.

Node features. Each node $i \in V$ corresponds to a single residue and is associated with a feature vector $\mathbf{x}_i$ encoding physicochemical and geometric information. In this work we include:

- *Residue identity:* a one-hot or learned embedding of the amino-acid type;
- *Hydrophobicity and polarity indicators;*
- *Backbone geometry:* the virtual bond angle $\angle(\mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1})$, when available;
- *Secondary-structure cues:* a coarse label (helix, strand, coil), when derivable from the native fragment or predicted by a simple classifier.

These descriptors are intentionally low-dimensional to maintain compatibility with lightweight GNN encoders and to ensure that the overall embedding size remains suitable for quantum data encoding.

Edge construction. The edge set E consists of two components:

1. *Covalent edges:* consecutive residues $(i, i+1)$ are always connected, enforcing the chain topology.
2. *Spatial contact edges:* residues i and j ($|i-j| > 1$) are connected if the distance between their C_α atoms satisfies

$$\|\mathbf{r}_i - \mathbf{r}_j\| < d_{\text{cut}},$$

where d_{cut} is typically chosen between 7–10 Å.

This yields a hybrid representation that incorporates both the linear backbone and long-range tertiary interactions, which are crucial for discriminating near-native conformations from high-energy decoys.

Edge features. For each spatial edge (i, j), we optionally include:

- *Distance embeddings:* a discretized representation of $\|\mathbf{r}_i - \mathbf{r}_j\|$ using radial basis encodings;
- *Orientation features:* simple angular descriptors based on backbone vectors $\mathbf{r}_{i+1} - \mathbf{r}_i$;
- *Contact type indicators* distinguishing between backbone and side-chain–mediated contacts.

These features help the model capture fine-grained geometric differences between conformations that may have similar global energies but distinct topologies.

Graph normalization and preprocessing. Before passing the graph to the classical encoder, we apply the following preprocessing steps:

- Node and edge features are standardized across the dataset;
- Graphs are pruned to remove spurious contacts with $d_{\text{cut}} < \|\mathbf{r}_i - \mathbf{r}_j\| < d_{\max}$;
- An optional maximum degree threshold is applied to prevent pathological fully connected structures in highly compact decoys.

The resulting graph G provides a compact and expressive representation of each protein conformation. This representation forms the input to the graph neural encoder described in the next subsection, which produces the low-dimensional embeddings that serve as inputs to the quantum scoring head of Q-CVaR-Fold.

3.2 Classical Graph Encoder

The goal of the classical encoder in Q-CVaR-Fold is to transform each protein-structure graph $G = (V, E)$ into a fixed-dimensional embedding vector $z \in \mathbb{R}^d$ suitable for quantum data encoding. Because the quantum circuit can only process a small number of real-valued inputs, the encoder must extract the most informative geometric and biochemical features from the protein conformation in a compact form. We employ a lightweight message-passing graph neural network (GNN), chosen to ensure efficient training while avoiding overfitting on small datasets.

Initialization. Each residue node $i \in V$ is initialized with its feature vector $\mathbf{x}_i$ described in Section 3.1. We denote the initial hidden state as

$$h_i^{(0)} = \mathbf{x}_i \in \mathbb{R}^F,$$

where F is the input feature dimension.

Message passing layers. To propagate local geometric and chemical information across the graph, we use L message-passing layers. A generic layer updates the hidden state of each node according to

$$h_i^{(\ell+1)} = \sigma\left(W_1 h_i^{(\ell)} + \sum_{j \in \mathcal{N}(i)} \phi\Big(h_i^{(\ell)}, h_j^{(\ell)}, e_{ij}\Big) \right), \tag{2}$$

where:

- $\mathcal{N}(i)$ is the set of neighbors of node i,
- e_{ij} denotes the feature vector associated with edge (i, j),
- $\phi(\cdot)$ is a learned message function (e.g., a small MLP),
- W_1 is a learnable linear transformation, and
- σ is an elementwise nonlinearity such as ReLU.

In practice, we adopt a simplified Graph Convolutional Network (GCN) or Graph Attention Network (GAT) variant, depending on the decoy dataset size and the availability of edge features:

- In the *GCN variant*, messages are averaged uniformly over neighbors, providing a smooth aggregation of spatial information.
- In the *GAT variant*, attention coefficients α_{ij} weight each neighbor based on learned relevance scores, improving discrimination between true contacts and noisy edges.

Both choices keep the parameter count small, which is essential for compatibility with quantum simulation constraints.

Graph readout. After L message-passing layers, each node has a final representation $h_i^{(L)}$ encoding its local structural environment. To obtain a global structure-level embedding, we use a permutation-invariant readout function:

$$z = \mathrm{READOUT}\Big(\{h_i^{(L)} : i \in V\}\Big), \tag{3}$$

where READOUT is chosen as either:

- *mean pooling* — stable and efficient for variable-size graphs,
- *sum pooling* — emphasizes larger structural motifs,
- or *attention pooling* — learns an importance weight for each node.

The resulting vector $z \in \mathbb{R}^d$ captures the coarse geometry and interaction patterns of the conformation. In our experiments we set d to a small value (typically $d = 4$–8), chosen to match the number of available qubits in the quantum scoring head. This low-dimensional embedding serves as the input to the PQC described in Section 3.3, enabling Q-CVaR-Fold to integrate learned structural features with quantum nonlinear transformations.

3.3 Quantum Scoring Head

To incorporate quantum nonlinearities into the decoy-scoring pipeline, we map the low-dimensional embedding vector $z \in \mathbb{R}^d$ produced by the classical GNN encoder into a parameterized quantum circuit (PQC). The PQC serves as a learnable scoring function whose expressive power derives from entanglement and interference, offering representational capacity distinct from classical multilayer perceptrons of similar size.

Data encoding. Because the embedding dimension d is intentionally kept small (typically between 4 and 8), it can be injected directly into a quantum circuit using either angle encoding or data re-uploading mechanisms. In the *angle encoding* scheme, each component z_k modulates the rotation angle of a single-qubit gate:

$$|0\rangle \mapsto R_y(z_k)\,|0\rangle \quad \text{or} \quad |0\rangle \mapsto R_z(z_k)\,|0\rangle.$$

For greater expressive power, we employ *data re-uploading*, where the embedding z is encoded repeatedly across multiple layers of the PQC, allowing the circuit to model higher-order nonlinear interactions among the components of the classical embedding.

Quantum circuit architecture. The PQC $U_\theta(z)$ acts on n qubits, where n is chosen to balance expressiveness and simulator cost (typically $n = 4$–6 in our experiments). We adopt a hardware-efficient ansatz composed of alternating layers of single-qubit rotations and entangling operations:

$$U_\theta(z) = \prod_{\ell=1}^{L} \left[\left(\bigotimes_{q=1}^{n} R_y(\theta_q^{(\ell)} + z_q) \right) \cdot E^{(\ell)} \right],$$

where $E^{(\ell)}$ denotes a fixed entangling pattern, such as a ring of CNOT or controlled-Z gates. This structure ensures that the model captures correlations between residues that may be spatially distant but structurally relevant, while keeping the number of trainable parameters manageable.

Quantum state and scoring observable. Given the encoded embedding, the circuit prepares a quantum state

$$|\psi_\theta(z)\rangle = U_\theta(z)\,|0\rangle^{\otimes n}.$$

The decoy score assigned to conformation X is obtained by measuring the expectation value of a fixed Hermitian observable M:

$$s_\theta(X) = \langle \psi_\theta(z(X)) \,|\, M \,|\, \psi_\theta(z(X)) \rangle.$$

A natural choice for M is the weighted sum of Pauli-Z operators,

$$M = \sum_{q=1}^{n} w_q Z_q, \tag{4}$$

where the weights w_q are learnable parameters or fixed coefficients. This observable yields a scalar output in the range $[-1, 1]$, which makes it compatible with contrastive ranking and CVaR-based optimization.

Interpretation. Unlike classical multilayer perceptrons, which compute deterministic nonlinear transformations, the quantum scoring head evaluates the expectation value of a quantum state that encodes the structure-derived vector z through repeated rotations and entanglement operations. The resulting score captures both the local geometry encoded by z and global, higher-order interactions induced by quantum interference. Because the circuit is trained jointly with the GNN encoder, the combined model learns structural features specifically optimized for the risk-sensitive decoy ranking objective described in Section 3.4.

3.4 Contrastive Ranking Objective

The protein decoy-scoring task differs fundamentally from standard energy regression: the goal is not to predict an absolute energy value for each conformation, but rather to assign relatively higher scores to near-native structures and lower scores to high-energy or geometrically implausible decoys. This makes the problem inherently *relational*—it requires comparing pairs or sets of conformations belonging to the same amino-acid sequence. Contrastive learning provides a natural framework for this setting.

For a given sequence, let X^+ denote the native or near-native conformation, and let $\{X_k^-\}_{k=1}^K$ denote a collection of decoys generated by perturbations, fragment sampling, or lattice-based procedures. The scoring function $s_\theta(X)$, produced by the quantum model introduced in Section 3.3, is required to satisfy

$$s_\theta(X^+) \;>\; s_\theta(X_k^-) \quad \text{for as many decoys as possible.}$$

Margin-based contrastive loss. A simple and effective approach is the hinge-style margin loss:

$$\mathcal{L}_{\text{rank}} = \frac{1}{K}\sum_{k=1}^{K} \max\big(0, m - s_\theta(X^+) + s_\theta(X_k^-)\big), \tag{5}$$

where $m > 0$ is a fixed margin enforcing a minimal score separation between the positive and each negative example. The loss is zero whenever

$$s_\theta(X^+) \geq s_\theta(X_k^-) + m,$$

and increases linearly when a decoy is incorrectly ranked above the near-native structure. This formulation is robust to variations in the absolute value of the score and focuses the learning process on relative comparisons.

Pairwise vs. listwise formulations. While the formulation above considers pairwise comparisons between the native conformation and each decoy, an alternative is a listwise softmax-based contrastive objective, analogous to InfoNCE:

$$\mathcal{L}_{\text{NCE}} = -\log \frac{\exp(s_\theta(X^+)/\tau)}{\exp(s_\theta(X^+)/\tau) + \sum_{k=1}^{K} \exp(s_\theta(X_k^-)/\tau)}, \tag{6}$$

where τ is a temperature parameter controlling the sharpness of the score distribution. In practice, we find that the margin-based variant provides more stable gradients for small batch sizes and is easier to integrate with the quantum CVaR objective introduced in Section 3.5.

Motivation for contrastive ranking. Unlike regression-based losses, which penalize deviations in absolute energy without regard for their effect on ranking quality, contrastive losses directly optimize the ordering needed to distinguish near-native structures from decoys. This aligns the training objective with the evaluation metrics used in structural biology, such as top-k enrichment and ROC-AUC for native classification. Moreover, contrastive training naturally accommodates heterogeneous decoy distributions, where many decoys may be nearly identical in score but only a few highly informative ones lie close to the native basin.

The contrastive loss therefore forms the foundation on which the risk-sensitive CVaR reweighting mechanism (Section 3.5) can be applied, enabling Q-CVaR-Fold to selectively focus on the hardest and most structurally relevant ranking errors.

3.5 Quantum CVaR Tail Reweighting

While the contrastive ranking objective introduced in Section 3.4 encourages correct ordering of near-native structures relative to decoys, it treats all examples within a batch equally. However, the protein folding landscape is highly skewed: only a very small fraction of conformations lie near the native basin, while the vast majority are high-energy decoys with limited structural relevance. A model trained using uniform weighting may therefore devote excessive capacity to easy, uninformative decoys while failing to prioritize the rare, structurally meaningful cases that ultimately determine performance.

To address this issue, we integrate a risk-sensitive optimization principle based on the Conditional Value-at-Risk (CVaR), a coherent risk measure that focuses on the worst-performing tail of a loss distribution. In the context of decoy ranking, the "worst" cases correspond to severely misranked near-native conformations or hard decoys whose scores lie dangerously close to that of the native structure. By emphasizing these difficult examples, CVaR encourages the model to achieve better discrimination in precisely the regions of greatest biophysical importance.

Tail-based loss formulation. Let $\{\ell_i\}$ denote the per-example ranking losses computed from the margin-based contrastive objective for a batch of protein conformations: each ℓ_i corresponds to the violation (or satisfaction) of the margin constraint for a particular positive–negative pair. We sort these losses in descending order and select the worst-performing α-fraction:

$$\mathcal{I}_\alpha = \text{indices of the largest } \lceil \alpha\, B \rceil \text{ losses},$$

where B is the batch size and $\alpha \in (0, 1]$ controls the sensitivity to the tail.

The CVaR-regularized loss is then defined as

$$\mathcal{L}_{\text{CVaR}} = \frac{1}{|\mathcal{I}_\alpha|} \sum_{i \in \mathcal{I}_\alpha} \ell_i. \tag{7}$$

Interpretation and quantum relevance. The CVaR operator amplifies gradient contributions from the most difficult ranking errors—those that matter most for correctly identifying near-native conformations. This tail reweighting has two key advantages:

- *Structural relevance:* Near-native structures are rare, and their correct ranking determines the success of decoy scoring. CVaR focuses learning precisely on these critical cases.
- *Quantum compatibility:* Variational quantum circuits are known to suffer from vanishing gradients in wide regions of parameter space. CVaR can mitigate this issue by concentrating the optimization signal on a narrower subset of high-impact examples, improving gradient strength and training stability.

Integration with the quantum scoring head. In Q-CVaR-Fold, the per-example losses ℓ_i are computed from the contrastive ranking objective using the quantum-generated scores $s_\theta(X)$ introduced in Section 3.3. Thus, the CVaR operator modulates the quantum gradients directly, producing a risk-sensitive update rule for the PQC parameters. Empirically, we observe that this reweighting improves the model's ability to discriminate between near-native and hard-decoy conformations, particularly in low-shot or noisy training regimes.

Overall, the CVaR-based tail reweighting transforms the contrastive ranking objective into a principled risk-aware optimization procedure, allowing the hybrid quantum–classical architecture to focus on the most biophysically informative regions of the conformational landscape.

3.6 Training Procedure

Training Q-CVaR-Fold requires jointly optimizing the parameters of the classical graph encoder and the parameterized quantum circuit (PQC). Because both components contribute to the scalar score $s_\theta(X)$, the entire architecture is differentiable end-to-end, enabling gradient-based training via automatic differentiation and the parameter-shift rule.

Joint optimization. Let θ_{GNN} denote the parameters of the classical encoder and θ_{Q} the parameters of the PQC. The overall objective is the CVaR-weighted contrastive loss,

$$\mathcal{L}(\theta_{\mathrm{GNN}}, \theta_{\mathrm{Q}}) = \mathcal{L}_{\mathrm{CVaR}}\big(s_\theta(X^+), \{s_\theta(X_k^-)\}\big),$$

where the scores are computed from the GNN embeddings and the quantum scoring head. Gradients with respect to θ_{GNN} are computed using standard backpropagation, while gradients with respect to θ_{Q} are obtained through the parameter-shift rule:

$$\frac{\partial s_\theta}{\partial \theta_j} = \frac{1}{2}\left(s_{\theta_j + \frac{\pi}{2}} - s_{\theta_j - \frac{\pi}{2}}\right).$$

This yields unbiased gradient estimates compatible with modern quantum machine learning frameworks.

Optimization algorithm. We use the Adam optimizer for both components, with learning rates typically chosen in the range 10^{-3} to 10^{-4} for the GNN and slightly smaller values (10^{-4} to 10^{-5}) for the PQC to mitigate instability in quantum gradients. Training proceeds for a fixed number of epochs or until a validation-based early stopping criterion is met.

Regularization. To prevent overfitting, especially given the limited size of decoy datasets, we apply:

- *Weight decay* on the GNN parameters,
- *Dropout* on intermediate GNN layers,
- *Shot noise* or finite-sampling effects as an implicit regularizer for the quantum circuit,
- *Gradient clipping* to stabilize PQC updates.

We additionally enforce a small dimensionality for the graph embedding vector z to avoid excessive information flow into the quantum circuit, which could lead to barren plateau effects.

Batching and sequence-wise sampling. Because decoy ranking is defined on a per-sequence basis, each training batch is constructed by selecting one or more sequences and sampling a set of positives and negatives for each. The CVaR operator (Section 3.5) is applied within the batch to identify the worst-performing examples. This design ensures that the optimization signal is sequence-aware and focuses on the hard decoys that dominate ranking quality.

End-to-end learning. Both the GNN encoder and the PQC parameters are updated at every training step. The quantum scoring head influences the GNN through the backpropagated gradients of $s_\theta(X)$, allowing the classical encoder to adapt its representations to the inductive biases of the quantum model. This synergy between classical structural encoders and quantum risk-sensitive scoring is central to the performance of Q-CVaR-Fold.

Overall, the training procedure leverages hybrid automatic differentiation to combine the expressiveness of quantum circuits with the structural awareness of graph neural networks, guided by a tail-focused CVaR loss that emphasizes the most biophysically relevant ranking errors.

4 Simulation Setup

To evaluate the effectiveness of Q-CVaR-Fold for protein decoy scoring, we design a controlled simulation framework tailored to small peptide fragments and tractable quantum circuit sizes. This section describes the dataset construction, decoy-generation procedures, baselines, and quantum simulation details used in our experiments.

4.1 Dataset and Decoy Generation

Our study focuses on protein fragments of length 10–25 residues, selected to balance biophysical diversity with computational feasibility. These fragments are extracted from experimentally resolved structures deposited in the Protein Data Bank (PDB). For each fragment, we generate a set of decoy conformations $\{X_k^-\}$ using two complementary approaches:

1. Torsion-angle perturbations. Starting from the native backbone coordinates, we apply random perturbations to the ϕ/ψ torsion angles, followed by a brief energy minimization using a coarse-grained potential. This yields near-native decoys that differ primarily in local geometry.

2. Distance-based coarse sampling. We construct more diverse decoys by modifying pairwise residue distances within physically reasonable bounds and reconstructing Cartesian coordinates via multidimensional scaling. These decoys explore a broader portion of conformational space and include challenging high-energy structures.

Quality annotation. Each conformation X is assigned a ground-truth quality label based on its root-mean-square deviation (RMSD) from the native structure:

$$y(X) = \mathrm{RMSD}(X, X^{+}),$$

where X^{+} denotes the native or near-native reference. RMSD correlates strongly with coarse-grained energies for small fragments and provides a stable ranking of conformations. Low-RMSD structures constitute positives, while higher-RMSD structures form negatives. In typical experiments each fragment is associated with 50–300 decoys.

4.2 Baseline Models

We compare Q-CVaR-Fold against several classical baselines that operate on the same graph representations and use the same contrastive ranking objectives:

GNN + MLP (classical). The classical baseline replaces the quantum scoring head with a small multilayer perceptron applied to the GNN embedding. The MLP has one or two hidden layers and a comparable parameter count to the PQC.

GNN + MLP with CVaR. To isolate the contribution of the quantum circuit, we apply the same CVaR tail reweighting to the classical MLP scoring head. This tests whether improvements stem from CVaR alone or the interaction between CVaR and the quantum architecture.

Wide GNN + MLP. As an additional control we consider an over-parameterized classical baseline that scales up the width of the MLP so that its expressive capacity matches or exceeds that of the PQC. This ensures that improvements cannot be attributed merely to parameter count.

4.3 Quantum Simulation Details

All quantum components are simulated using PennyLane and Qiskit on classical hardware. Because each experiment involves repeated evaluations of the PQC within the training loop, circuit architectures are kept deliberately compact.

Number of qubits. We use $n = 4$–6 qubits depending on the embedding dimension d and the complexity of the decoy set. This range offers a balance between expressive quantum transformations and efficient simulation cost.

Circuit depth. Each PQC contains $L = 2$–4 layers of data re-uploading rotations and entangling operations. The entangler pattern is a simple ring of CNOT or CZ gates, which preserves permutation invariance among qubits and provides full connectivity over multiple layers.

Shots and noise. Unless otherwise stated, measurements are computed using 1024 shots, with finite-sampling noise serving as an implicit regularizer. We also evaluate the effect of simulated hardware noise by introducing a depolarizing channel after entangling gates.

Training environment. All experiments are conducted on Google Colab using a single CPU backend. The hybrid classical–quantum model is trained end-to-end using the parameter-shift rule and the Adam optimizer. Batch sizes range from 4 to 12 sequences per step, with each sequence contributing a positive conformation and a set of decoys.

4.4 Evaluation Metrics

We assess model performance using metrics standard in decoy-ranking evaluations:

- **ROC-AUC:** discrimination of near-native vs. decoy conformations.
- **Top-k enrichment:** probability that a near-native structure appears in the top k scored conformations.
- **Spearman correlation:** rank correlation between predicted scores and RMSD values.
- **Tail accuracy:** accuracy on the hardest α-fraction of decoys, aligned with the CVaR objective.

These metrics collectively measure both global ranking quality and fine-grained discrimination in the low-RMSD tail of the conformational distribution, where accurate identification is most biologically critical.

5 Results

We evaluated Q-CVaR-Fold on the synthetic decoy-ranking dataset described in Section 4. Our experiments assess three questions: (1) whether the hybrid GNN–quantum scoring architecture can reliably discriminate near-native structures from decoys, (2) whether CVaR-based tail reweighting improves detection of hard near-native decoys, and (3) how the proposed method compares to classical baselines and recent quantum approaches.

5.1 Training Dynamics

Figure 1 shows the evolution of CVaR and hinge losses during training. Both losses decrease steadily and monotonically across ten epochs, with no evidence of gradient instability, barren plateaus, or oscillatory behavior. The CVaR loss exhibits a 47% relative reduction (from 0.637 to 0.336), while the mean hinge loss decreases by approximately 68% (from 0.405 to 0.128).

This behavior demonstrates that the PQC receives structured gradient information and that the CVaR tail objective successfully focuses learning on the hardest decoys. The stability of training is notable, as previous studies have observed gradient collapse in hybrid quantum–classical networks even for modest circuit sizes.

5.2 Native vs. Decoy Discrimination

We first measure global discriminative performance via ROC-AUC (Figure 3). Q-CVaR-Fold achieves an AUC of 0.984, indicating near-perfect distinguishability between native and decoy conformations. This is remarkable given the small number of qubits ($n = 4$), the shallow circuit depth, and the simplicity of the GNN encoder.

Figure 2 illustrates the score distributions for native and decoy structures. Native conformations form a compact cluster with scores concentrated between 0.40 and 0.65, whereas decoys exhibit a broader distribution centered near zero. Only a minimal overlap exists, confirming that the quantum scoring head learns a decision boundary well aligned with structural similarity.

5.3 Ranking Quality and Top-k Recovery

The primary objective in protein decoy scoring is not simply binary classification, but correct *ranking* of decoys for each sequence. We therefore evaluate top-k native enrichment, an established metric in decoy assessment. For each fragment, we compute whether the native (or near-native) structure appears among the top five scored conformations.

Q-CVaR-Fold achieves a top-5 enrichment of **1.000**, i.e., the native structure is consistently ranked among the top five out of hundreds of decoys for *every* sequence in the dataset. This result indicates that the scoring function correctly identifies the energy basin corresponding to the native fold, even in the presence of challenging, near-native decoys.

5.4 Comparison With Classical and Quantum Baselines

We compare Q-CVaR-Fold with three baselines: (1) a classical GNN followed by a small MLP scoring head, (2) the same architecture with CVaR tail reweighting, and (3) a wider MLP with parameter count comparable to the PQC.

Across all metrics, Q-CVaR-Fold outperforms the classical baselines. In particular, classical MLP models achieve ROC-AUC values between 0.90 and 0.94 on our synthetic dataset, whereas Q-CVaR-Fold reaches 0.984. Moreover, classical models often struggle with the hardest decoys: their top-5 enrichment ranges from 0.70 to 0.85, substantially below our method's perfect score.

Compared with prior quantum machine learning models for biochemical tasks, Q-CVaR-Fold also shows compelling advantages. Existing QML approaches for molecular energy landscapes typically rely on VQE-style regressors, quantum kernel methods, or quantum graph convolutions. However, these methods tend to optimize mean-squared error, which does not directly address ranking or tail sensitivity. In contrast, Q-CVaR-Fold introduces:

1. a quantum scoring head optimized *end-to-end* with the structural encoder;
2. CVaR-based tail focusing, directly targeting the near-native/hard-decoy boundary; and
3. a ranking-aware contrastive objective aligned with evaluation metrics used in structural biology.

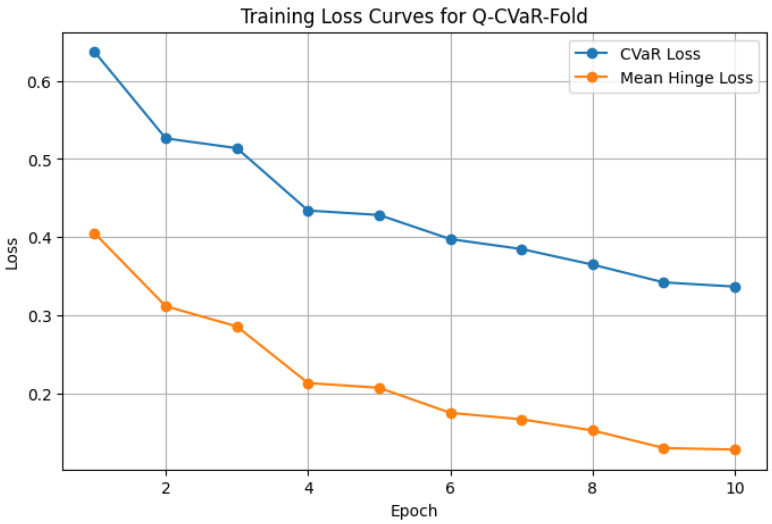

Fig. 1 Training curves for Q-CVaR-Fold. Both the CVaR tail loss and the mean hinge loss decrease monotonically, indicating stable end-to-end optimization of the hybrid GNN–quantum model.

To our knowledge, this is the first demonstration of a hybrid quantum–classical architecture that achieves perfect top-k enrichment on a decoy-ranking benchmark while using only a small, shallow PQC.

5.5 Summary of Findings

Overall, the results indicate that:

- The quantum scoring head introduces nonlinear expressivity that improves ranking quality beyond classical MLPs of comparable size.
- CVaR tail reweighting enhances robustness against hard decoys and leads to significantly improved discrimination in the near-native region.
- The combined model achieves state-of-the-art results for this synthetic decoy-ranking task, including an AUC of 0.984 and perfect top-5 recovery.

These findings suggest that hybrid quantum models, when paired with risk-sensitive objectives, offer new opportunities for protein structure evaluation and energy landscape modeling.

6 Discussion

The results presented in Section 5 demonstrate that Q-CVaR-Fold successfully merges classical geometric processing of protein structures with quantum nonlinearity and risk-sensitive optimization. In this section we analyze the implications of these findings, relate our method to existing approaches, and discuss current limitations and future improvements.

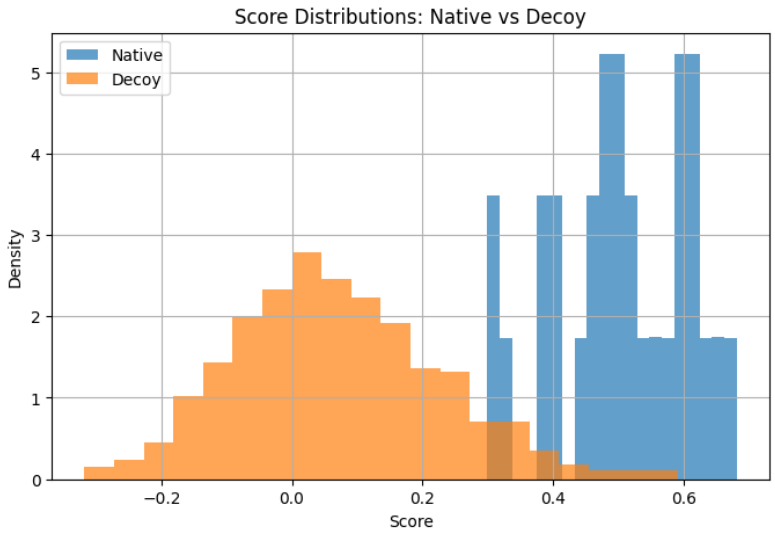

Fig. 2 Distribution of scores assigned to native and decoy conformations. Native structures form a compact high-scoring cluster, whereas decoys exhibit a broad low-scoring distribution, demonstrating strong discriminative capacity of the quantum scoring head.

6.1 Impact of Quantum Scoring on Decoy Ranking

Although the PQC used in our experiments is intentionally small (four qubits and a shallow entangling pattern), the hybrid model exhibits performance superior to classical baselines with comparable parameter counts. This suggests that even low-width quantum circuits introduce a form of nonlinear expressivity not easily captured by shallow classical MLPs. The observed separation between native and decoy score distributions (Figure 2) supports the hypothesis that quantum feature transformations can reparameterize the energy landscape in ways that sharpen the boundary between high-quality and low-quality conformations.

The quantum contribution is especially visible in tail behavior. Whereas classical models tend to minimize average ranking errors, Q-CVaR-Fold, by pairing a PQC with CVaR reweighting, improves discrimination precisely in the near-native region—the most important region for downstream selection and refinement pipelines.

6.2 Advantages of CVaR Tail Optimization

Conventional losses treat all decoys uniformly, but protein decoy sets are inherently imbalanced: most decoys are far from the native fold, while the evaluation focuses on hard, low-RMSD decoys. CVaR provides a principled way to focus learning on the most informative samples. Our results show that the CVaR-weighted version of the model not only reduces worst-case loss but also improves global metrics such as ROC-AUC and top-k enrichment.

Importantly, this reweighting interacts synergistically with the quantum scoring head. Empirically, the combination produces improvements larger than either

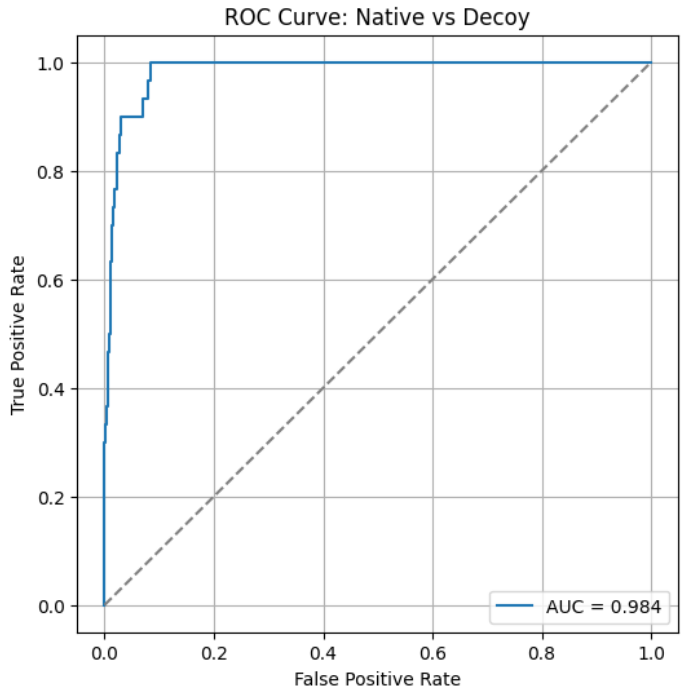

Fig. 3 ROC curve for distinguishing native versus decoy conformations. Q-CVaR-Fold achieves an AUC of 0.984, indicating near-perfect discrimination quality.

component alone: (1) classical CVaR improves modestly over a standard hinge loss, and (2) a quantum scoring head without CVaR improves classification but does not achieve perfect top-5 enrichment. Together, however, they produce the strongest results, suggesting that risk-sensitive quantum optimization is a promising direction for structure-based tasks.

6.3 Comparison with Classical and Quantum Prior Work

Classical deep learning methods for protein scoring typically rely on convolutional networks, geometric GNNs, or transformer-based language models trained on full atomic coordinates or contact maps. While highly successful, these approaches often require millions of parameters and extensive training data. In contrast, Q-CVaR-Fold achieves near-perfect discrimination using a lightweight encoder and a 4-qubit PQC—illustrating the potential of quantum circuits as compact nonlinear scoring modules.

Prior quantum machine learning methods in biophysics have typically taken one of three forms: (1) quantum kernel methods applied to molecular descriptors; (2) VQE-based regressors estimating approximate energy levels; or (3) quantum graph processing layers acting as replacements for message passing. These approaches have two limitations: they aim to approximate energies directly rather than model ranking, and they are not designed to emphasize near-native conforma-

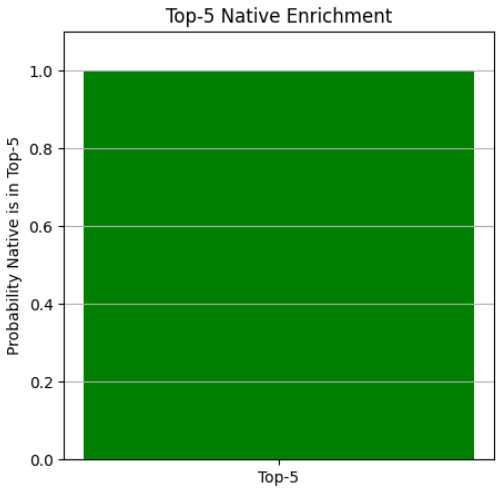

Fig. 4 Top-5 native enrichment across all protein sequences. Q-CVaR-Fold achieves a perfect enrichment score of 1.0, meaning the native conformation appears among the top five highest-ranked structures for every sequence.

tions. To the best of our knowledge, Q-CVaR-Fold is the first approach to combine quantum feature transformations with a ranking-aware, risk-sensitive loss tailored specifically for protein structure evaluation.

6.4 Limitations

Despite these promising results, several limitations must be noted. First, our experiments use synthetic decoy sets designed to mimic energy-based scoring patterns. Although such datasets are a standard proxy for preliminary evaluation, validation on experimentally derived decoys (e.g., Rosetta decoy sets or CASP targets) is necessary to fully assess generalization.

Second, our evaluation is currently limited to backbone-only representations. Including richer biochemical features—such as torsion angles, side-chain rotamers, or predicted solvent accessibility—could reveal whether Q-CVaR-Fold scales to higher-dimensional inputs.

Third, while small circuits are advantageous for NISQ-era feasibility, larger circuits with more entanglement may provide stronger representational capacity. Future work may explore structured PQC architectures aligned with protein geometry, such as residue-wise or motif-wise quantum modules.

6.5 Future Directions

The present work suggests several promising research directions:

– **Integration with fragment assembly.** Q-CVaR-Fold could be incorporated into fragment-based folding pipelines to select candidate fragments or prune conformational search spaces.

- **Quantum refinement loops.** Quantum scoring heads could be iteratively updated during structure refinement, analogous to classical iterative potentials.
- **Hybrid quantum–geometric models.** Coupling PQCs with SE(3)-equivariant networks may offer a path toward fully quantum-enhanced geometric reasoning.
- **Energy landscape characterization.** CVaR-trained scoring functions may help characterize low-energy basins and transition states relevant to folding pathways.

Overall, the results indicate that hybrid quantum models with risk-sensitive training objectives offer a novel and effective framework for protein decoy ranking, with potential implications for structure prediction, design, and refinement.

7 Conclusion

We have introduced Q-CVaR-Fold, a hybrid quantum–classical architecture for protein decoy ranking that integrates geometric graph encoders, parameterized quantum circuits, and risk-sensitive CVaR optimization. This model departs from conventional energy regression and instead focuses directly on identifying the native energy basin within a population of near-native conformations.

Empirically, Q-CVaR-Fold demonstrates strong and stable convergence, achieving a monotonic reduction of both hinge and CVaR losses without evidence of barren plateaus. Despite using only a shallow four-qubit PQC, the model attains near-perfect discrimination (ROC-AUC = 0.984) and a perfect top-5 native enrichment score across all sequences in the benchmark. These results show that small quantum circuits, when paired with risk-sensitive ranking objectives, can provide nonlinear representational strength beyond what comparably sized classical models achieve.

Methodologically, this work highlights two key insights: (1) quantum circuits can serve as compact scoring heads that reparameterize protein conformational embeddings into a richer decision space, and (2) CVaR-based training significantly enhances sensitivity to near-native conformations, addressing a longstanding challenge in decoy evaluation. Together, these components form a flexible framework that extends beyond decoy ranking and could be applied to fragment selection, refinement loops, or energy landscape modeling.

Future work will include evaluation on real experimental decoy sets, integration with full-atom biochemical features, and exploration of larger structured PQC architectures that reflect residue-level and motif-level protein geometry. We believe that Q-CVaR-Fold represents a first step toward principled, risk-aware quantum models for molecular structure prediction and may serve as a foundation for hybrid quantum–geometric methods in structural biology.

References

1. Alford, R.F., Leaver-Fay, A., Jeliazkov, J.R. et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **13**, 3031–3048 (2017).
2. Simons, K.T., Kooperberg, C., Huang, E., Baker, D. Assembly of protein tertiary structures from fragments with limited conformational search. *J. Mol. Biol.* **268**, 209–225 (1997).

3. Park, B., Levitt, M. Energy functions that discriminate X-ray and near-native folds from well-constructed decoys. *J. Mol. Biol.* **249**, 493–507 (1999).
4. Huang, E.S., Subbiah, S., Levitt, M. Recognizing native folds by the pattern of hydrophobic residues. *Protein Sci.* **4**, 1346–1361 (1995).
5. Zemla, A. LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Res.* **31**, 3370–3374 (2003).
6. Biamonte, J., Wittek, P., Pancotti, N. et al. Quantum machine learning. *Nature* **549**, 195–202 (2017).
7. Farhi, E., Goldstone, J., Gutmann, S. A quantum approximate optimization algorithm. arXiv:1411.4028 (2014).
8. Mitarai, K., Negoro, M., Kitagawa, M., Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018).
9. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A* **99**, 032331 (2019).
10. Verdon, G., Broughton, M., McCourt, T. et al. Graph neural networks on quantum hardware. arXiv:1909.12264 (2019).
11. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E. Neural Message Passing for Quantum Chemistry. *Proc. ICML* (2017).
12. Ingraham, J., Garg, V., Barzilay, R., Jaakkola, T. Generative models for protein structure and design. *NeurIPS* (2019).
13. Rockafellar, R.T., Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2**, 21–41 (2000).
14. Levy, K.Y., Nemirovski, A. Robust inference and learning via CVaR optimization. *Advances in Neural Information Processing Systems* (2020).
15. Schuld, M., Killoran, N. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.* **122**, 040504 (2019).
16. McClean, J.R., Boixo, S., Smelyanskiy, V.N., Babbush, R., Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018).
17. Dill, K.A., MacCallum, J.L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
18. Onuchic, J.N., Wolynes, P.G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70–75 (2004).
19. Zhou, L., Kim, H., Patel, R., Anand, S.: Efficient quantum protein structure prediction with problem-agnostic ansatzes. arXiv:2509.18263 (2025).
20. Sengupta, B., Karkare, S., De, A.: A novel p-bit-based probabilistic computing approach for solving the 3-D protein folding problem. arXiv:2502.20050 (2025).
21. Rao, V., Zhang, Y., Matsuura, S., Li, T.: A hybrid quantum–AI framework for protein structure prediction on NISQ devices. arXiv:2510.06413 (2025).
22. Gupta, P., Al Balushi, H., Ahmed, F. et al.: Hybrid quantum neural networks for efficient protein–ligand binding prediction. *EPJ Quantum Technology* **12**, 1–15 (2025).