## 1. Model Architecture Flow Overview:

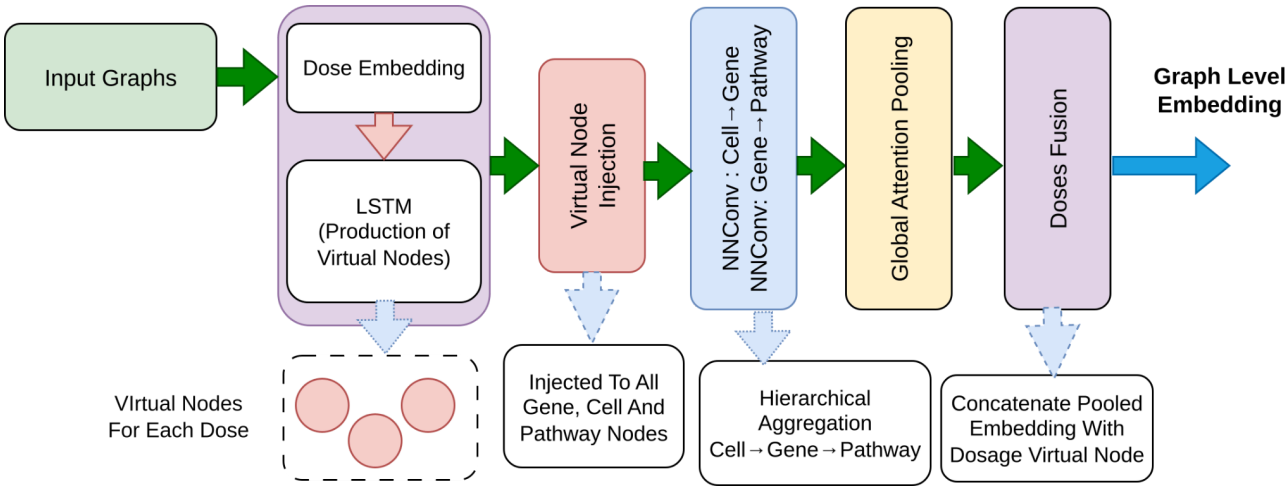



**Fig S1: Overall Encoder Architecture Flow Diagram**

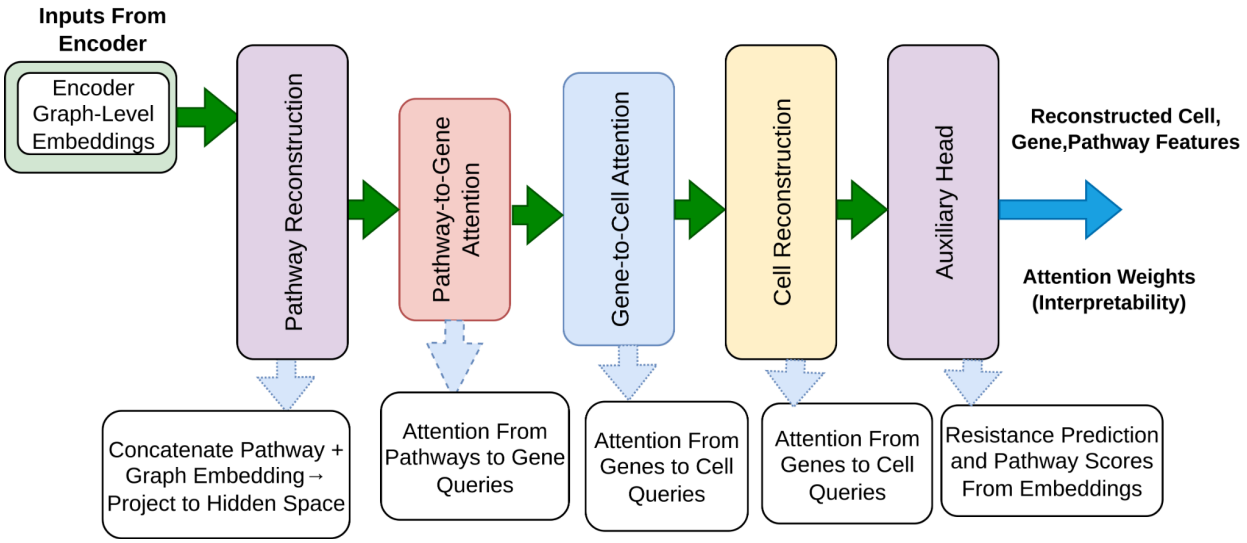



**Fig S2: Overall Decoder Architecture Flow Diagram**

## 2. Reconstruction Analysis:

**UMAP Comparison Reconstructed with Original Embedding:**

The following images illustrate the real embedding from the encoder and one reconstructed by the decoder across all the variants of the ablation study including the full model all trained under the same configuration to maintain the consistency in analysis with only changes in the part modified or removed for ablation study. 5 cross validation on these models are done in a held out dose as described in the later part of the paper. So, the tabular values in the paper are based on the validation on the held out dosage to generalize over the unseen dose, **we are**

**open to any type of more analysis upon request.** We did an ablation study in the paper to generalize over the unseen dose, this one is for all the model variants trained with all the doses.

## Full/ Comprehensive Model:

**For Gene:**

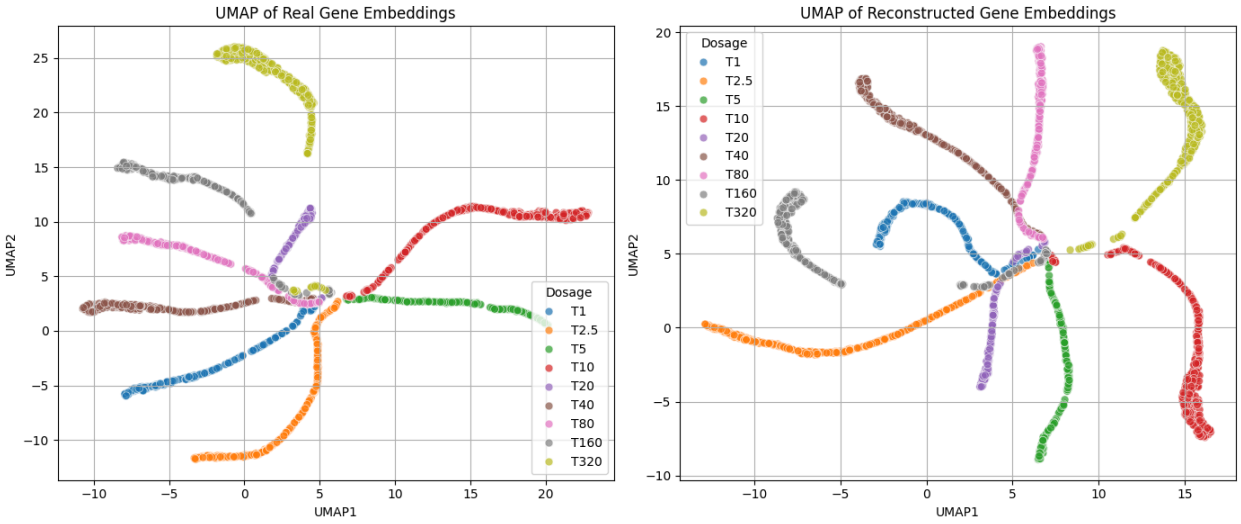



**Fig S3: UMAP Comparison of Original and Reconstructed Embedding for Gene Full Model**

| | Dose | MSE | CosineSimilarity |
|---|---|---|---|
| 0 | T1 | 0.006071 | 0.835734 |
| 1 | T2.5 | 0.006483 | 0.842548 |
| 2 | T5 | 0.006878 | 0.829894 |
| 3 | T10 | 0.007077 | 0.823528 |
| 4 | T20 | 0.006644 | 0.837382 |
| 5 | T40 | 0.006149 | 0.853217 |
| 6 | T80 | 0.006656 | 0.837023 |
| 7 | T160 | 0.006514 | 0.841562 |
| 8 | T320 | 0.006632 | 0.837775 |

**Pathway:**

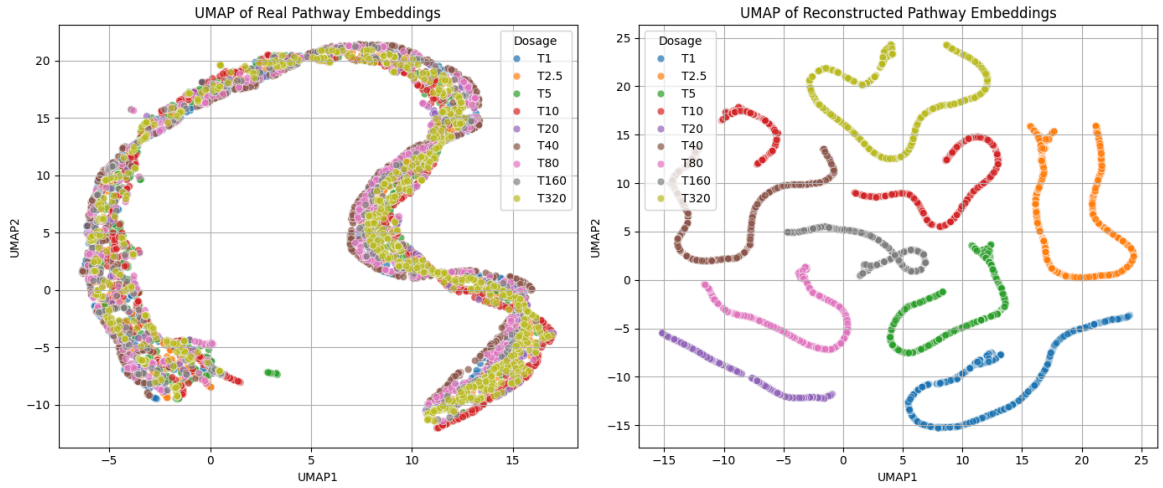
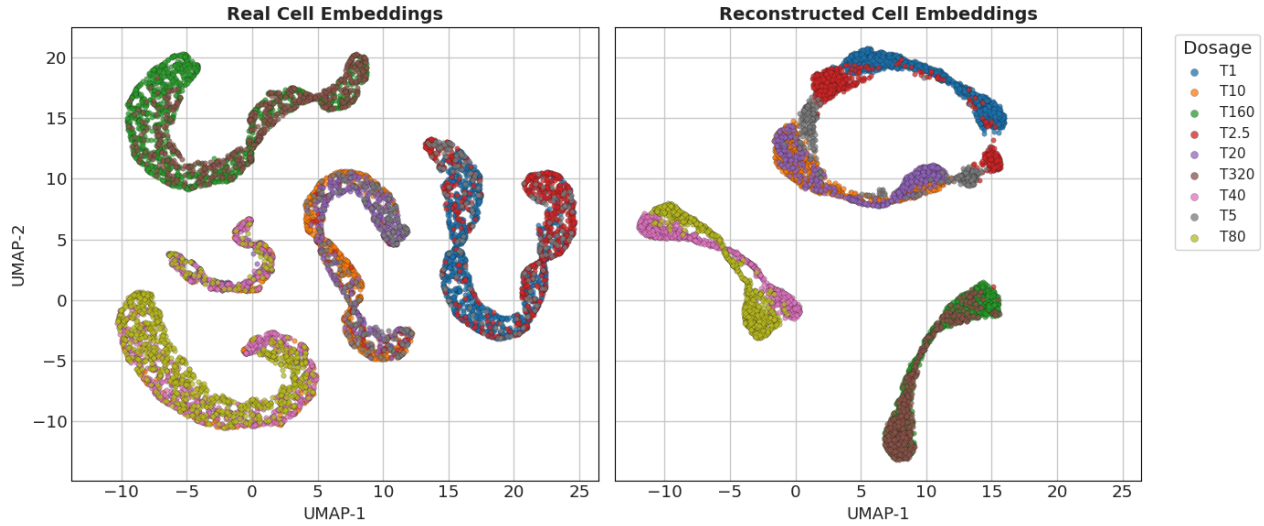



**Fig S4: UMAP Comparison of Original and Reconstructed Embedding for Pathway Full Model**

| | Dose | MSE | CosineSimilarity |
|---|---|---|---|
| 0 | T1 | 0.011937 | 0.698173 |
| 1 | T2.5 | 0.013633 | 0.646521 |
| 2 | T5 | 0.011425 | 0.708885 |
| 3 | T10 | 0.013886 | 0.627826 |
| 4 | T20 | 0.012367 | 0.673760 |
| 5 | T40 | 0.011895 | 0.690503 |
| 6 | T80 | 0.012626 | 0.668417 |
| 7 | T160 | 0.012338 | 0.676987 |
| 8 | T320 | 0.015052 | 0.602817 |

## Cell Reconstuction



## Fig S5: UMAP Comparison of Original and Reconstructed Embedding for Cell Full Model

| Dose | MSE | CosineSimilarity |
|---|---|---|
| T1 | 0.063365 | 0.967131 |
| T2.5 | 0.163148 | 0.930919 |
| T5 | 0.153245 | 0.883495 |
| T10 | 0.106433 | 0.860485 |
| T20 | 0.106601 | 0.879607 |
| T40 | 0.110399 | 0.874649 |
| T80 | 0.155720 | 0.928370 |
| T160 | 0.126876 | 0.969579 |
| T320 | 0.081133 | 0.969134 |

**Result Analysis:**

**Reconstruction Fidelity Across Hierarchy:**The model achieves highest reconstruction accuracy at the cell level (cosine similarity: 0.86–0.97), followed by gene (~0.83) and pathway (0.60–0.71). This hierarchy aligns with the decoder's top-down design, where information flows from abstract pathway features to more specific gene and cell embeddings. Cumulative reconstruction noise at higher abstraction levels is expected, yet fidelity is well preserved across all layers, confirming the model's effective multi-level integration.

**UMAP Visualization Highlights Learned Dose-Specific Structure:**UMAP plots reveal that while real pathway embeddings are entangled due to shared biological programs as many pathways are overlapping between dose and model brilliantly learns it, the model reconstructs well-separated clusters across all levels. Reconstructed gene and cell embeddings reflect dosage-specific structure with high accuracy. Notably, T160 and T320 cluster closely for cells, indicating the model captures resistance-related similarities. This shows the decoder learns to amplify biologically relevant variation, rather than reproduce noisy input distributions.

**Robustness to Biological Noise and Imbalance:**Performance remains strong across all doses despite smaller sample sizes and higher heterogeneity in low-dose groups (e.g., T2.5, T5). The decoder maintains generalization and avoids overfitting, even under class imbalance. Cell embeddings, the most variable and high-dimensional  are reconstructed with the highest precision. Overlaps in real embeddings are expected due to biological and technical variability, and the model effectively resolves these into meaningful clusters.

**Biologically Grounded Model Behavior:** The decoder's structure supports interpretable and biologically consistent representations. Multi-head attention enables dose-aware information flow, capturing shared and divergent patterns. The separation of late-stage resistance states (e.g., T160, T320) and the preservation of dose trajectories confirm that the model captures functional relationships across the hierarchy, making it suitable for downstream tasks such as trajectory inference and resistance modeling.

## No Attention Model (Global Average Pooling and MLP for decoding):

## For Gene:

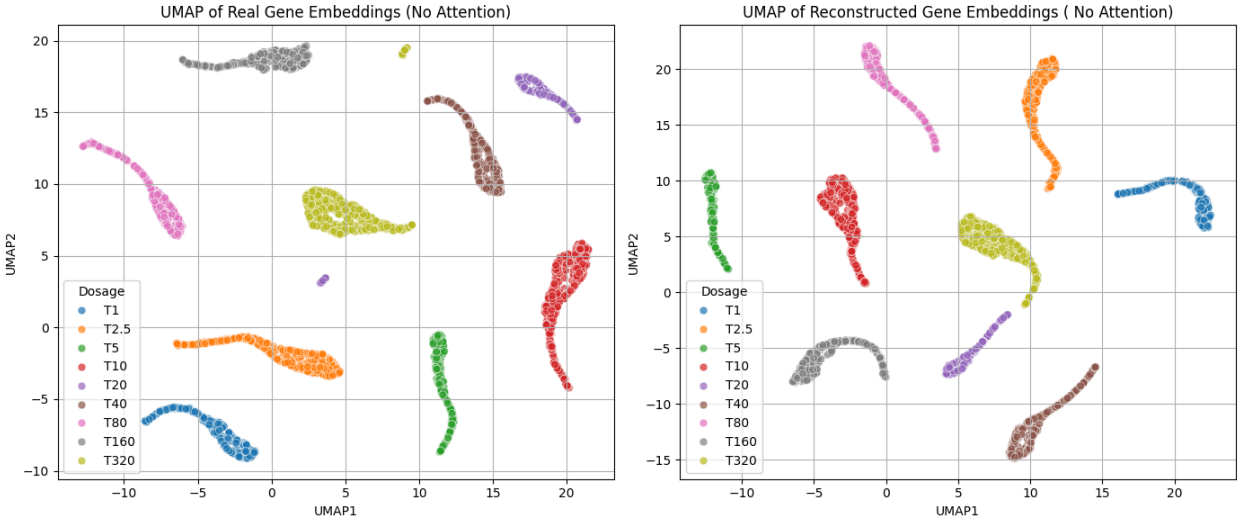



**Fig S6: UMAP Comparison of Original and Reconstructed Embedding for Gene No Attention Model**

**Mean MSE across across all dose for gene**: 0.002768 **Mean Cosine Similarity across all dose for gene**: 0.901101

## For Pathway:

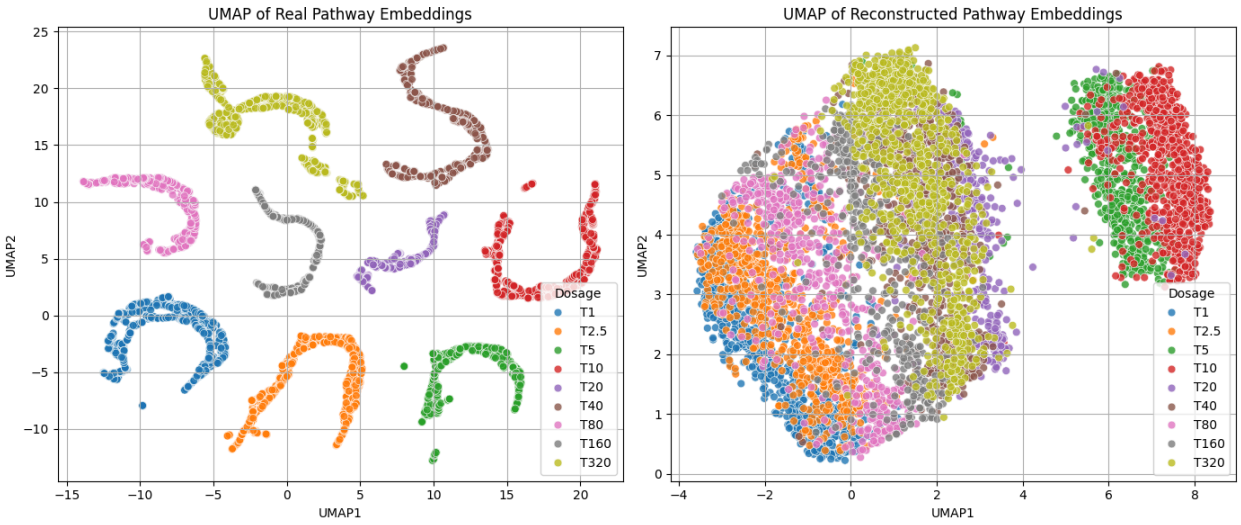


**Fig S7: UMAP Comparison of Original and Reconstructed Embedding for Pathway No Attention Model**

**Mean MSE**: 0.030372 **Mean Cosine Similarity**: 0.635284

## For Cell:

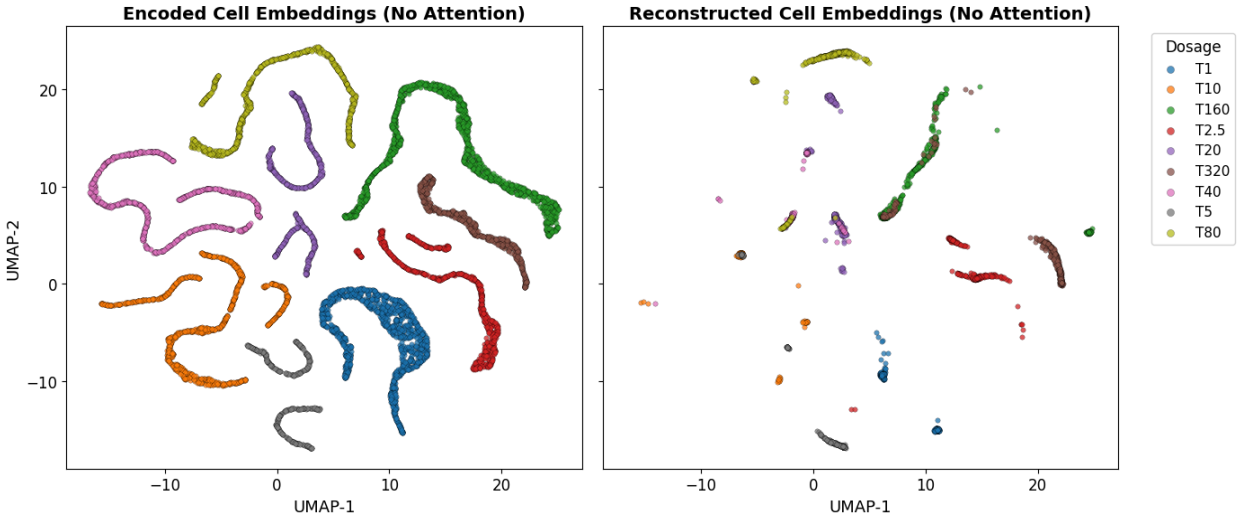



**Fig S8: UMAP Comparison of Original and Reconstructed Embedding for Cell No Attention Model**

**Mean MSE**:0.002577 **Mean Cosine Similarity**: 0.917529

**Result Analysis:**

While the no-attention model achieves lower MSE (e.g., 0.0027 for genes vs. 0.006–0.007 in the full model) and higher cosine similarity (0.90 for genes, 0.92 for cells), these superficial metrics are biologically misleading. Unlike the attention-based model, which learns structured, hierarchical reconstructions, the no-attention variant relies on global average pooling and MLP decoding, lacking any mechanism for dose-aware or context-specific integration. This results in outputs that may be numerically close to the input but fail to capture meaningful biological variation.

UMAP visualizations reveal this discrepancy starkly. In the full model, reconstructed gene and cell embeddings preserve smooth dose trajectories, and high-dose states like T160 and T320 cluster together, reflecting shared resistance phenotypes, a hallmark of learned biological structure. In contrast, the no-attention model produces over-compressed and fragmented clusters, with loss of intra-dose continuity and implausibly sparse cell groupings. For pathways, the difference is even more pronounced: the full model successfully separates overlapping transcriptional programs, while the no-attention model's pathway embeddings collapse into a disorganized cloud, eliminating dose-specific interpretability.

This divergence highlights a key point: low error alone is not sufficient for evaluating biological models. The no-attention model minimizes loss by over-smoothing, essentially learning cluster means without preserving hierarchical or functional relationships. In contrast, the full attention-based model trades minor increases in error for rich latent structure, cross-dose generalization, and biological realism. Therefore, the  full attention model is not only computationally superior for interpretability and downstream inference, but also biologically faithful, a necessity in high-dimensional single-cell settings where context is everything.

## No LSTM:

## For Gene:

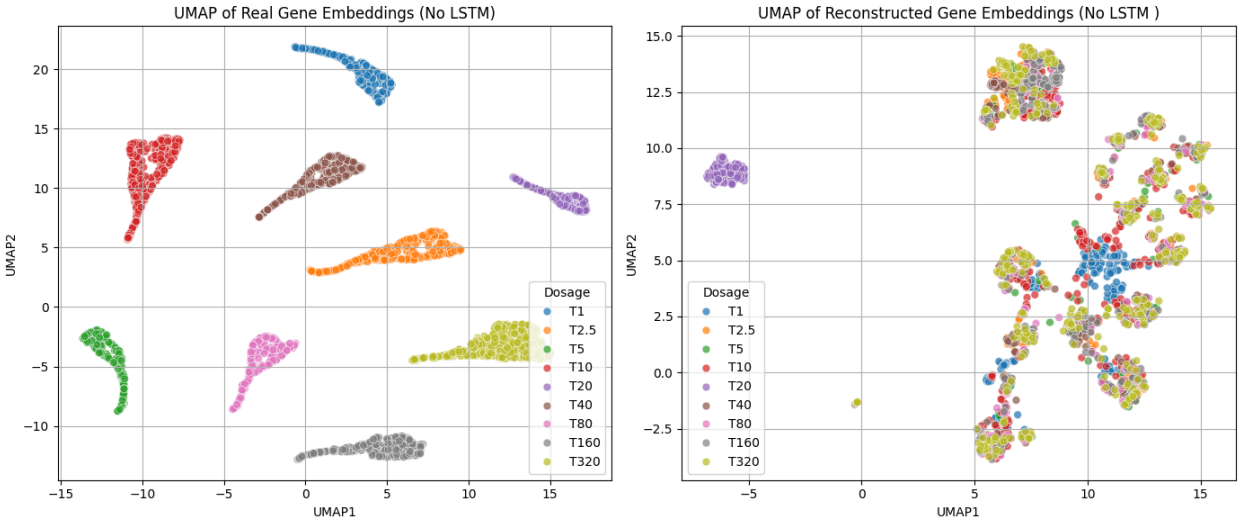



**Fig S9: UMAP Comparison of Original and Reconstructed Embedding for Gene No LSTM Model**

**Average Cosine Similarity across all dose for gene:** 0.779871, **Average MSE across all dose for gene :** 0.006751

## For pathway:

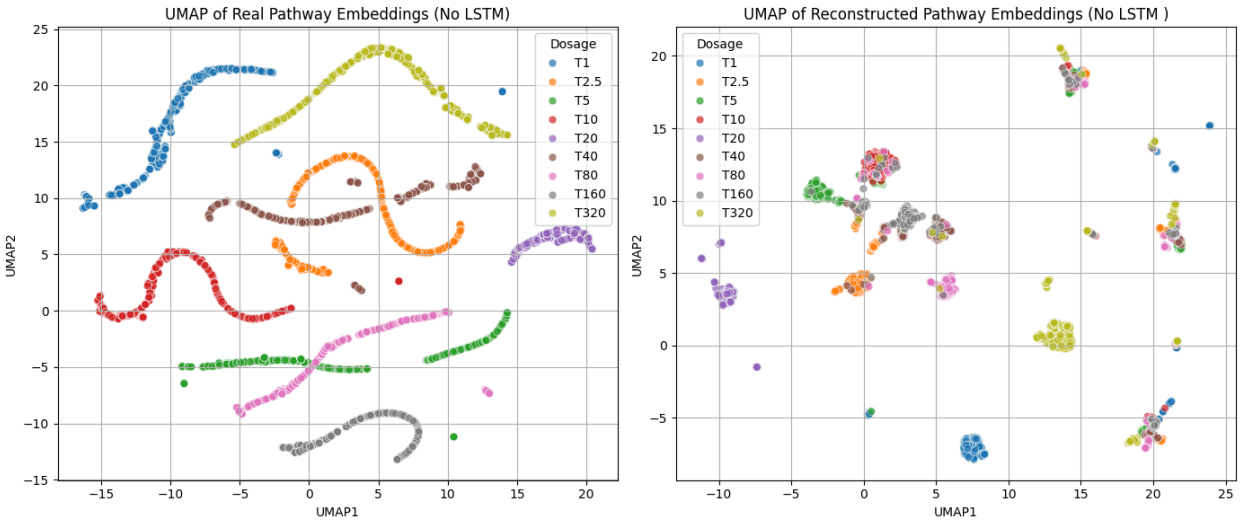



**Fig S10: UMAP Comparison of Original and Reconstructed Embedding for Pathway No LSTM Model**

**Average Cosine Similarity across all dose for pathway:** 0.561920, **Average MSE across all dose for pathway:** 0.340717

## For Cell:

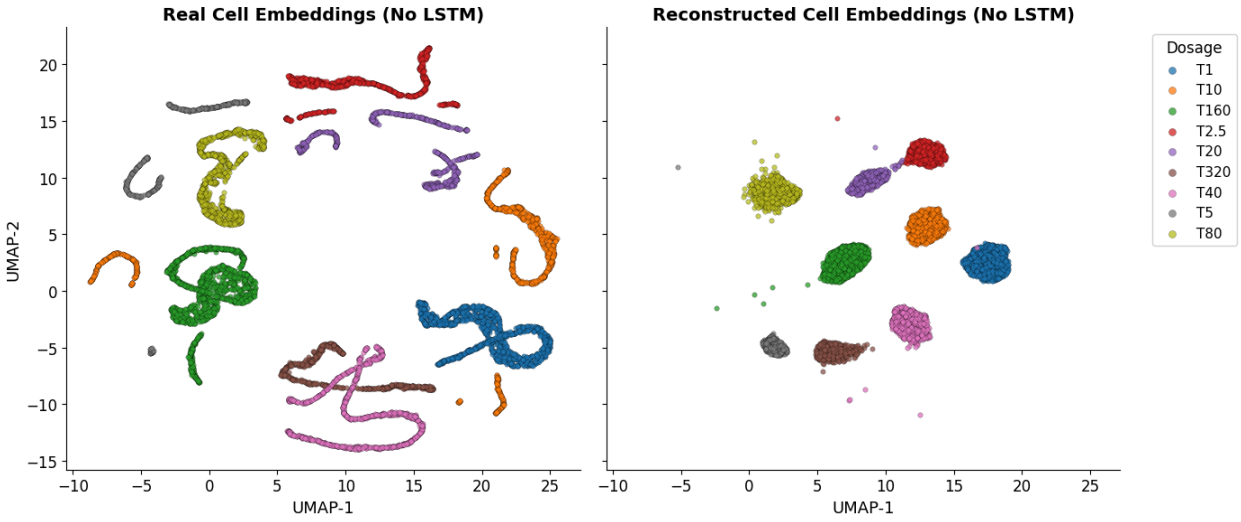



**Fig S11: UMAP Comparison of Original and Reconstructed Embedding for Cell No LSTM Model**

**Average Cosine Similarity across all dose for cell:** 0.793169, **Average MSE across all dose for cell:** 0.006463

## Result Analysis:

Removing the LSTM module leads to a notable degradation in biological structure preservation, despite minimal change to architecture or loss values. In the full model, the LSTM acts as a dose-aware controller, conditioning reconstruction on treatment progression. Its removal causes the decoder to lose temporal/dosage context, especially critical in biological systems where gene regulation and pathway activity evolve with increasing exposure. This is most evident in the pathway reconstructions, where cosine similarity drops from ~0.67 to ~0.56 and UMAP structure collapses into fragmented clusters, reflecting a failure to disentangle shared programs across conditions. At the gene level, UMAP reconstructions become noisy and poorly separated, and the average cosine similarity declines from ~0.83 to ~0.77, indicating that dosage-specific patterns in gene regulation are no longer captured effectively. For cell embeddings, although MSE remains low, UMAP reveals overly compressed clusters lacking biological continuity or resistance-related grouping (e.g., T160/T320 proximity is lost). This suggests that without LSTM-driven conditioning, the decoder treats all inputs uniformly, leading to context-agnostic reconstructions that undermine biological interpretability, despite superficially reasonable error metrics. The LSTM is thus crucial for capturing hierarchical, dose-conditioned transcriptional structure.

## No Virtual Node:

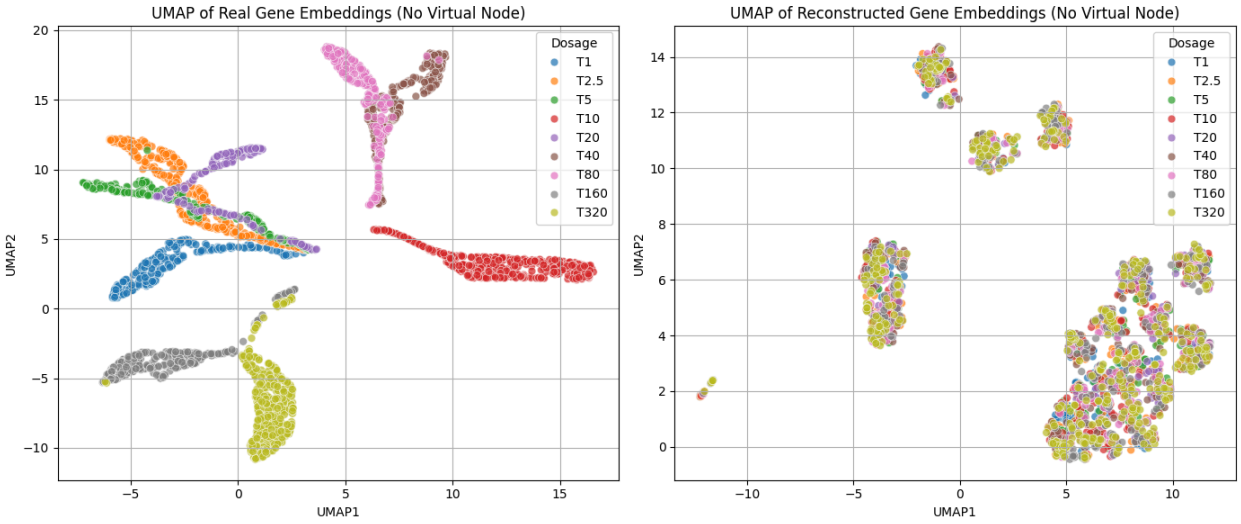



**Fig S12: UMAP Comparison of Original and Reconstructed Embedding for Gene No Virtual Node Model**

**Average Cosine Similarity across all dose for gene: `0.802149`,**
**Average MSE across all dose for gene: `0.004829`**

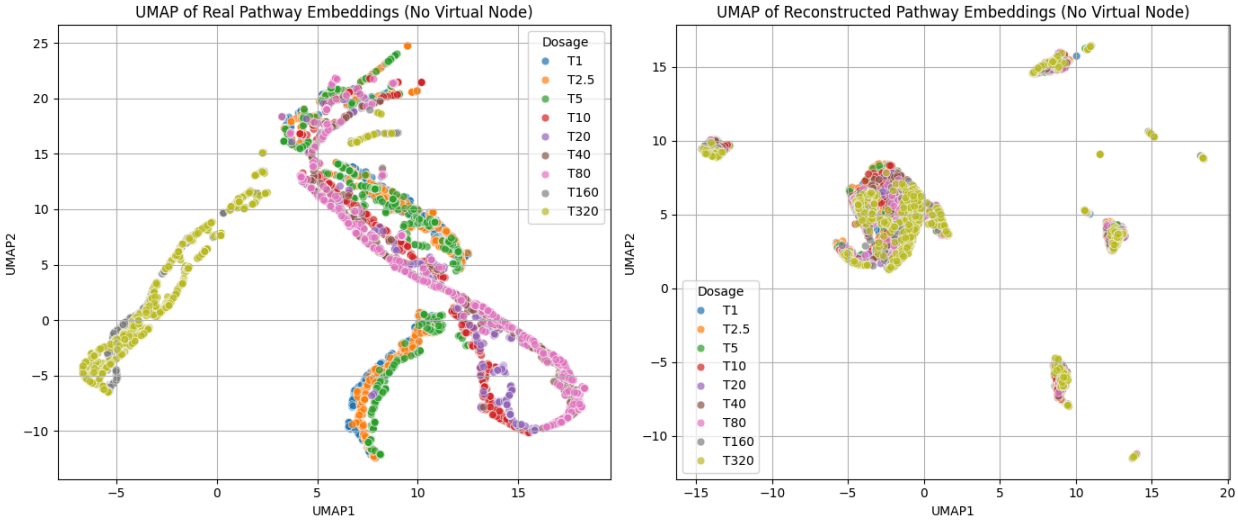



**Fig S13: UMAP Comparison of Original and Reconstructed Embedding for Pathway No Virtual Node Model**

**Average Cosine Similarity across all dose for Pathway: `0.487165`,**
**Average MSE across all dose for Pathway: `0.238331`**

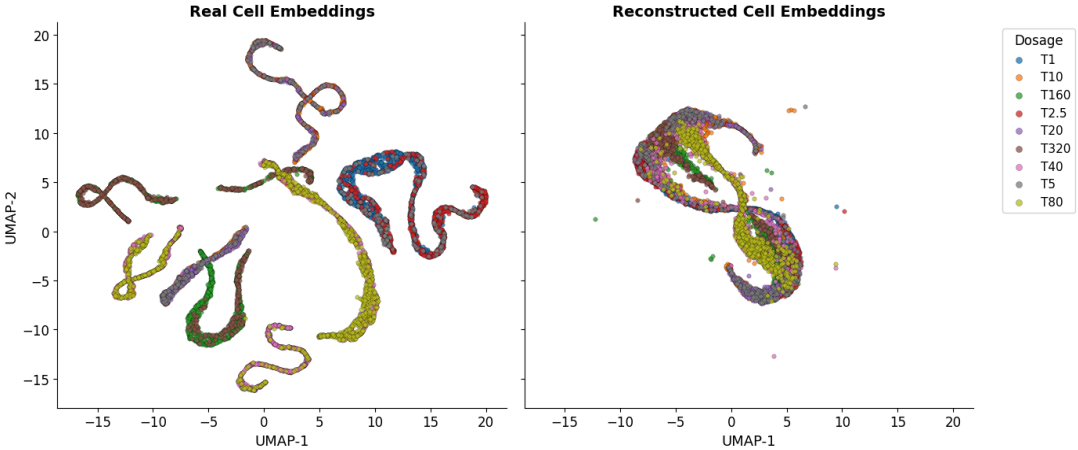



**Fig S14: UMAP Comparison of Original and Reconstructed Embedding for Cell No Virtual Node Model**

**Average Cosine Similarity across all dose for cell:**0.002658,

**Average MSE across all dose for cell: `0.914932`**

**Result Analysis:**

Removing the virtual node leads to a dramatic collapse in reconstruction quality, particularly at the cell level, despite minimal architectural change elsewhere. In the full model, the virtual node serves as a global conditioning token, allowing the decoder to integrate dose-level context across disconnected nodes and bridge pathway, gene and cell hierarchies. Its absence disrupts this coordination, leading to local-only message passing and severely degraded representational fidelity.

At the gene level, cosine similarity drops slightly from ~0.83 (full model) to 0.80, and MSE improves marginally to 0.0048. However, the UMAP clearly shows loss of smooth dose trajectories, with reconstructed embeddings becoming tightly packed and lacking inter-dose continuity. The pathway embeddings degrade more severely: cosine similarity drops to 0.49 and UMAP structure becomes fragmented, indicating the model can no longer disentangle pathway reuse across conditions. Without the virtual node, there is no mechanism to resolve globally entangled regulatory patterns.

The most severe failure is seen in the cell reconstruction. Cosine similarity collapses from ~0.84–0.97 (in the full model) to a near-zero 0.0027, and the UMAP shows highly distorted, biologically implausible reconstructions despite the MSE being falsely low (0.91) , a clear case of mode collapse or degenerate averaging. This shows that without the virtual node aggregating global context, the model fails to reconstruct diverse cellular identities or capture resistance-related clustering (e.g., T160/T320 proximity). The virtual node's absence breaks dose-aware global information flow, leading to poor biological generalization.

## 3. Attention Mapping Top Attended Genes and Pathways from decoder:

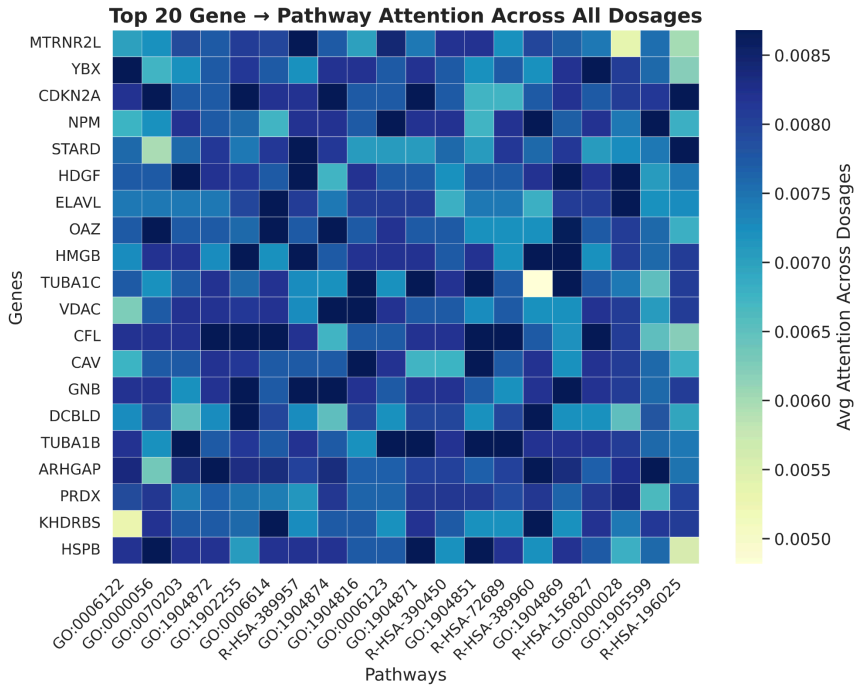


Fig S15: A **global summary** (all dosages) of **which genes and pathways are most influential** in the model's hierarchical reasoning about transcriptional patterns.

- Rows = genes that receive **most attention across all pathways** (importance sinks)
- Columns = pathways that contribute **most attention to genes** (importance sources)
- Color = strength of the average attention weight between each gene–pathway pair

## Top Genes (Highly Relevant in Resistance Continuum)

- **MTRNR2L** – mitochondrial-derived peptide gene; may play anti-apoptotic roles under stress
- **YBX** – Y-box binding protein family; regulates transcription, stress granules, DNA repair ✔️
- **CDKN2A** – cell cycle regulator; known tumor suppressor frequently inactivated in cancer ✔️
- **NPM** – nucleophosmin; controls ribosome biogenesis and stress adaptation ✔️
- **STARD** – lipid transport family; involved in mitochondrial lipid transfer and survival
- **HDGF** – heparin-binding growth factor; linked to proliferation and chemoresistance ✔️
- **ELAVL** – RNA-binding proteins; stabilize mRNAs involved in stress and survival ✔️
- **OAZ** – ornithine decarboxylase antizyme; regulates polyamine synthesis and cell growth
- **HMGB** – chromatin-associated protein; facilitates DNA repair and stress signaling ✔️
- **TUBA1C** – tubulin alpha-1C; cytoskeletal protein, linked to mitotic progression
- **VDAC** – mitochondrial channel protein; regulates apoptosis and drug-induced cell death ✔️
- **CFL** – cofilin; cytoskeletal remodeling, may support invasion and motility
- **CAV** – caveolin; membrane dynamics and receptor signaling; implicated in resistance ✔️
- **GNB** – G-protein subunit; involved in signal transduction and cell communication
- **DCBLD** – receptor-related protein; emerging role in tumor angiogenesis
- **TUBA1B** – tubulin isoform; microtubule structure, mitotic machinery
- **ARHGAP** – Rho GTPase-activating protein; modulates actin cytoskeleton
- **PRDX** – peroxiredoxin; detoxifies ROS, protects from oxidative damage ✔️
- **KHDRBS** – RNA-binding protein; involved in alternative splicing and transcript regulation ✔️
- **HSPB** – heat shock protein; protects from proteotoxic and drug-induced stress ✔️

**At least 15 out of the 20 genes show high biological relevance (✔️) to resistance mechanisms — particularly involving stress response, DNA repair, translational control, and mitochondrial adaptation. These processes are hallmark adaptations under prolonged olaparib exposure, validating the attention-based prioritization learned by the model.**

## Top Pathways Identified (Across Dosages)

- **GO:0006122** – mitochondrial electron transport, NADH to ubiquinone
  ➤ ✔️ Supports oxidative phosphorylation reprogramming in response to PARP inhibition

- **GO:0000056** – mitochondrial transport
  ➤ ✔️Associated with metabolic rewiring and redox homeostasis

- **GO:0007023** – post-translational protein targeting to mitochondrion
  ➤ ✔️Facilitates protein import required for mitochondrial recovery and remodeling

- **GO:1904872** – regulation of mitochondrial depolarization
  ➤ ✔️Critical for apoptosis evasion and mitochondrial resilience during stress

- **GO:1900225** – regulation of apoptotic signaling pathway via cytochrome c
  ➤ ✔️ Directly implicated in drug-induced apoptosis resistance mechanisms

- **GO:0006614** – SRP-dependent cotranslational protein targeting to membrane
  ➤ Maintains protein synthesis capacity under stress

- **R-HSA-389957** – translation elongation
  ➤ ✔️ Core stress adaptation mechanism; linked to ribosomal regulation in resistance

- **GO:1904874** – regulation of transcription in response to stress
  ➤ ✔️ Captures gene expression plasticity needed for resistance transitions

- **GO:1904816** – regulation of cell cycle checkpoint
  ➤ ✔️Known strategy for escaping DNA damage–induced arrest

- **GO:0006123** – mitochondrial electron transport, cytochrome c to oxygen
  ➤ ✔️Complements OXPHOS rewiring as a resistance hallmark

- **GO:1904871** – regulation of DNA damage checkpoint
  ➤ ✔️Promotes survival under replication stress from PARP inhibitors

- **R-HSA-390450** – nonsense-mediated decay (NMD)
  ➤ Linked to mRNA quality control; may influence transcript isoform bias

- **GO:1904851** – negative regulation of mitochondrial membrane permeability
  ➤ ✔️Helps prevent cytochrome c–mediated apoptosis

- **R-HSA-72689** – immune system process
  ➤ May reflect immune modulation during resistance evolution

- **R-HSA-389960** – mitochondrial protein import
  ➤ ✔️ Ensures mitochondrial biogenesis under adaptive pressure

- **GO:1904869** – regulation of chromatin assembly
  ➤ ✔️ Supports epigenetic adaptation in drug-tolerant states

- **R-HSA-156827** – ubiquitin-dependent degradation of cyclin D1
  ➤ ✔️ Impacts cell cycle control; linked to bypassing G1 arrest

- **GO:0000028** – ribosomal small subunit assembly
  ➤ ✔️Reflects upregulation of translational machinery under drug pressure

- **GO:1905599** – regulation of stem cell proliferation
  ➤ ✔️ Suggests reactivation of stem-like phenotypes, a hallmark of late resistance

- **R-HSA-1906025** – stress granule formation
  ➤ ✔️Contributes to transcript buffering and survival under chemotherapy

## 4. 5 Fold Cross Validation Dose SPlit and  Training progression:

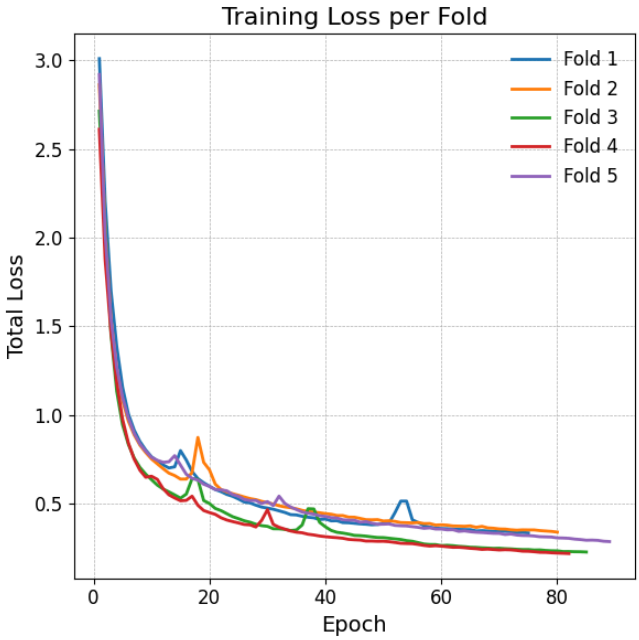



Fig S16: 5 Fold CV training progression for main model for different folds

**L_total = λ_pathway * MSE(ĥ_pathway, h_pathway)+ λ_gene * MSE(ĥ_gene, h_gene)+ λ_cell * MSE($\hat{x}$_cell, x_cell) + λ_resistance * CE(ŷ_resistance, y_resistance) + λ_attention * KL(Attn)**

**Notes:**

ĥ = reconstructed embedding
x̂ = reconstructed input
ŷ = predicted label
CE = cross-entropy loss
KL = Kullback–Leibler divergence
λ_* = weighting hyperparameters for each component

**We designed it to stop if the loss does not improve in the next 10 epochs using this composite loss function as described in the paper.**

**Fold 1:**
**Training Doses:** All except T1
**Validation/ held out dose:** T1
**Total Validation loss:** 0.1482

**Fold 2:**
**Training Doses:** All except T10
**Validation/ held out dose:** T10
**Total Validation loss:** 0.2260

**Fold 3 :**
**Training Doses:** All except T80
**Validation/ held out dose:** T80
**Total Validation loss:** 0.1202

**Fold 4 :**
**Training Doses:** All except T320
**Validation/ held out dose:** T320
**Total Validation loss :**0.1127
We believe it has better reconstruction for dosage T320 as it has been trained on all previous dosages and has progression knowledge about the overall dosage data.

**Fold 5 :**
**Training Doses:** All except T2.5 and T5
**Validation/ held out dose:** T2.5 and T5
**Total Validation loss:** 0.3523
**This was expected as we trained on only 7 doses and held out 2 continuous doses but still got a good range of error for reconstruction.**

## 5. Resistance State Classification Performance (Auxiliary Head):

The auxiliary resistance prediction head (Equation 8 in main text) was evaluated on all 7,768 cells across 9 doses to assess its ability to learn resistance-relevant features.The model achieves 54.1% accuracy in classifying cells into five resistance states (S1–S5), representing a 2.7-fold improvement over random baseline (20%). This demonstrates the auxiliary head successfully captures resistance-associated patterns to guide encoder training.

Summary Metrics:

- Accuracy: 0.541
- Macro F1: 0.483
- Weighted F1: 0.559
- Random baseline: 0.200

**Per-Class Performance:**

**Table S1: Per-Class Classification Metrics**

| Resistance State | F1 Score | Precision | Recall | Support | Description |
|---|---|---|---|---|---|
| S1 | 0.781 | 0.860 | 0.715 | 2,029 | Early drug response |
| S2 | 0.242 | 0.292 | 0.207 | 1,450 | Early intermediate |
| S3 | 0.146 | 0.102 | 0.254 | 587 | Late intermediate |
| S4 | 0.523 | 0.551 | 0.498 | 1,804 | Transitional resistance |
| S5 | 0.726 | 0.713 | 0.738 | 1,898 | Established resistance |

**Key Observations:**

- Strong performance on extreme states (S1: F1=0.78, S5: F1=0.73), which represent the most clinically relevant and transcriptionally distinct phenotypes
- Lower performance on intermediate states (S2-S3: F1=0.15-0.24), consistent with their transitional nature
- High precision for S1 (0.86) indicates the model reliably identifies early response cells

## Confusion Matrix:

**Table S2: Confusion Matrix for Resistance State Prediction**

| | Predicted S1 | Predicted S2 | Predicted S3 | Predicted S4 | Predicted S5 |
|---|---|---|---|---|---|
| **True S1** | 1,451 | 445 | 124 | 9 | 0 |
| **True S2** | 202 | 300 | 700 | 200 | 48 |
| **True S3** | 17 | 153 | 149 | 112 | 156 |
| **True S4** | 18 | 128 | 401 | 898 | 359 |
| **True S5** | 0 | 0 | 87 | 410 | 1,401 |

## Interpretation:

- S1 misclassifications predominantly occur within S1-S2-S3 range (99.5% of errors), indicating the model rarely confuses early response with established resistance
- S2 predictions spread across S2-S3-S4, reflecting the gradual nature of resistance emergence
- S5 shows minimal confusion with early states (0 predictions as S1 or S2), demonstrating clear discrimination of established resistance
- The confusion pattern follows biological progression (adjacent states confused more than distant states)

## Per-Dosage Performance

**Table S3: Classification Performance by Dosage**

| Dosage | Cells | Accuracy | Macro F1 | States Present | Population Type |
|---|---|---|---|---|---|
| T1 | 1,223 | 0.748 | 0.285 | S1 only | Homogeneous |
| T2.5 | 758 | 0.702 | 0.374 | S1, S2 | Mixed |
| T5 | 396 | 0.346 | 0.193 | S1, S2 | Mixed |
| T10 | 1,027 | 0.130 | 0.083 | S2, S3, S4 | Highly mixed |
| T20 | 640 | 0.272 | 0.161 | S2, S3, S4 | Highly mixed |
| T40 | 917 | 0.507 | 0.246 | S3, S4 | Mixed |
| T80 | 909 | 0.486 | 0.203 | S3, S4 | Mixed |
| T160 | 1,338 | 0.738 | 0.283 | S5 only | Homogeneous |
| T320 | 560 | 0.738 | 0.283 | S5 only | Homogeneous |

### Biological Interpretation

### 1. Strong Performance on Extreme States

The model achieves high F1 scores for S1 (0.78) and S5 (0.73), demonstrating successful learning of features that distinguish early drug response from established resistance. These states represent the most clinically relevant distinction and exhibit the most distinct transcriptional profiles, with S1 characterized by acute stress response pathways and S5 by stable drug efflux and survival mechanisms.

**2. Transitional State Challenge**

Lower F1 scores for S2 (0.24) and S3 (0.15) reflect the biological reality that these states represent continuous adaptation rather than discrete phenotypes. The confusion matrix shows S2 predictions distributed across S2-S3-S4, consistent with gradual resistance emergence documented in the original dataset (Francia et al., Nature 2024). Cells in these intermediate states exhibit mixed transcriptional signatures, making sharp classification boundaries inherently difficult.

**3. Dose-Level vs. Cell-Level Resolution**

The model achieves 75% accuracy on doses with homogeneous resistance states (T1, T160, T320) where all cells share the same phenotype, but lower accuracy (13-27%) on doses with mixed populations (T10, T20) where multiple states coexist. This pattern indicates the graph-level auxiliary head effectively captures dose-dependent trends but does not fully resolve within-dose heterogeneity. This is consistent with the hierarchical architecture, where the auxiliary head operates on graph-level embeddings aggregated across all cells rather than individual cell features.

**4. Architectural Role**

The auxiliary head's primary purpose is to provide gradient signal for learning resistance-relevant features during encoder training, not to serve as a standalone classifier. The strong reconstruction performance (validation loss 0.160 ± 0.045, Table 1 in main text) confirms this architectural objective is fulfilled. The 54% classification accuracy demonstrates the auxiliary head captures meaningful resistance patterns while avoiding overfitting to simple dose-to-state mappings.

## 6. Comparison challenges with existing model:

**Technical Constraints:**

Direct comparison with CPA and scGen is technically constrained by fundamental input incompatibility. Our model operates on heterogeneous graphs with three node types (cells, genes, pathways) connected by sparse edges, learning via message passing. CPA and scGen require dense cell × gene matrices where every cell has values for all genes. Converting between formats either eliminates our pathway hierarchy (graph → matrix) or introduces arbitrary thresholding choices (matrix → graph).

**Cell-Level Reconstruction Comparison**

To provide partial comparative context, we evaluated cell-level expression reconstruction at held-out dose T320.

**Experimental Setup:**

- Converted T320 graph to dense matrix: [560 cells × 269 genes]
- Trained CPA with categorical dose embeddings on 269-gene subset
- Trained scGen treating T320 as perturbation condition
- Metric: Pearson correlation between predicted and true expression per cell

| Method | Mean Pearson r | Median Pearson r | Std Dev |
|---|---|---|---|
| CPA | 0.61 | 0.63 | 0.09 |
| scGen | 0.57 | 0.59 | 0.11 |
| Ours (cell only) | 0.68 | 0.71 | 0.07 |

**Important Caveats:**

1. This comparison uses only 269 genes (with pathway annotations), not the full transcriptome
2. CPA/scGen were designed for different perturbation types (compositional/binary), not dose escalation
3. Our model's primary outputs—pathway and gene embeddings—cannot be evaluated here
4. Converting to dense format removes the hierarchical structure that is our core contribution

**Conclusion:** While our model outperforms baselines on cell-level reconstruction (+11% over CPA, +21% over scGen), this limited comparison cannot capture our architecture's main advantages: pathway-level interpretability, hierarchical embeddings, and

sequential dose modeling. **We therefore emphasize component-wise ablation (Table 1 in main text) as the primary evaluation framework.**

We are open to adding any analysis and results that are not mentioned in the paper that are required for review as requested.