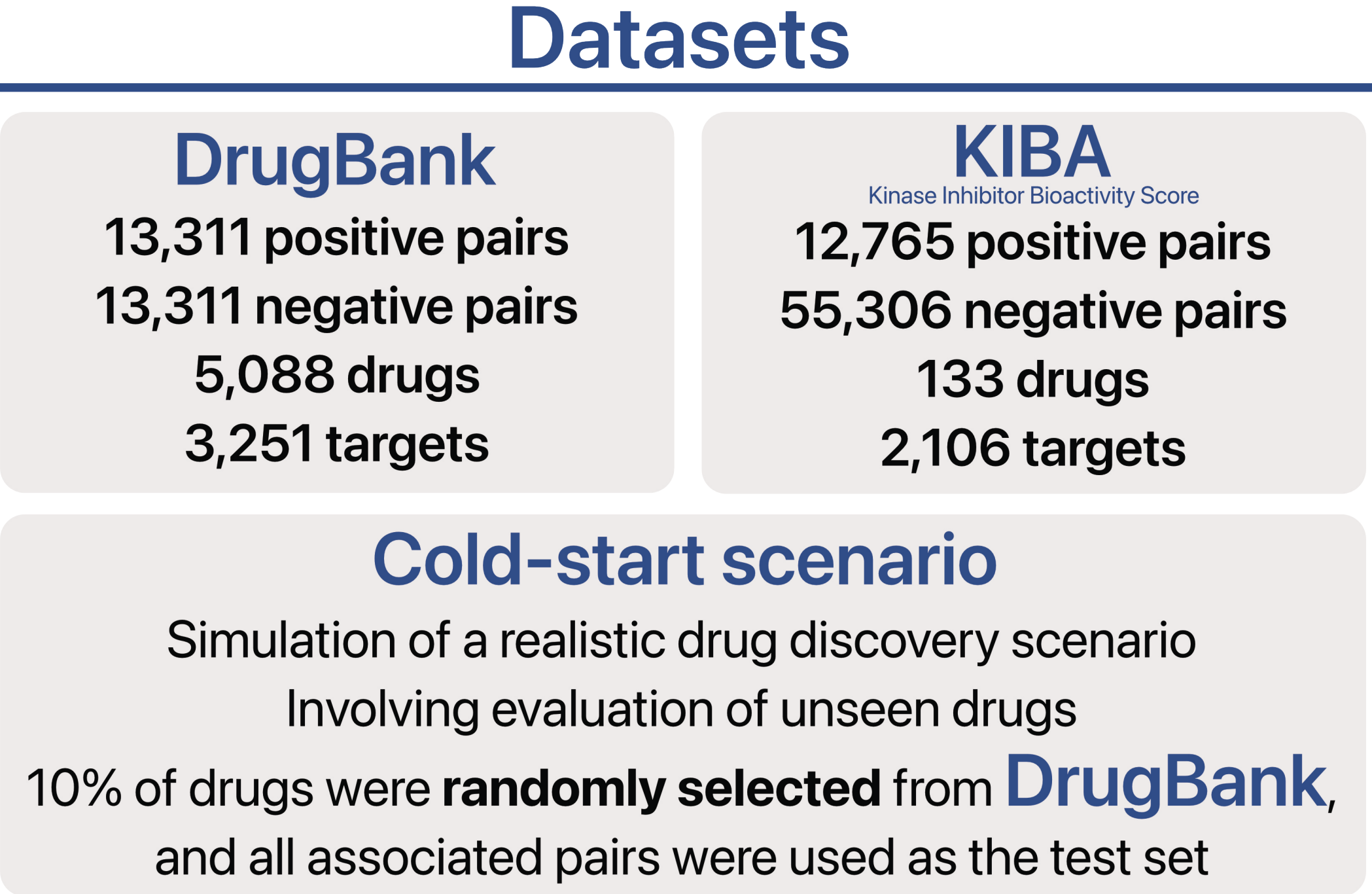


3D structure-aware embeddings

- **3D protein structures** fetched from PDB and AlphaFoldDB
- **3D drug conformers** generated by RDKit's ETKDG algorithm followed by MMFF94 force-field optimisation
- Frozen **massively pre-trained** encoders

Co-attention module

- Bidirectional cross-attention with feed-forward layers
- **Architectural consideration** for cross-modal interactions

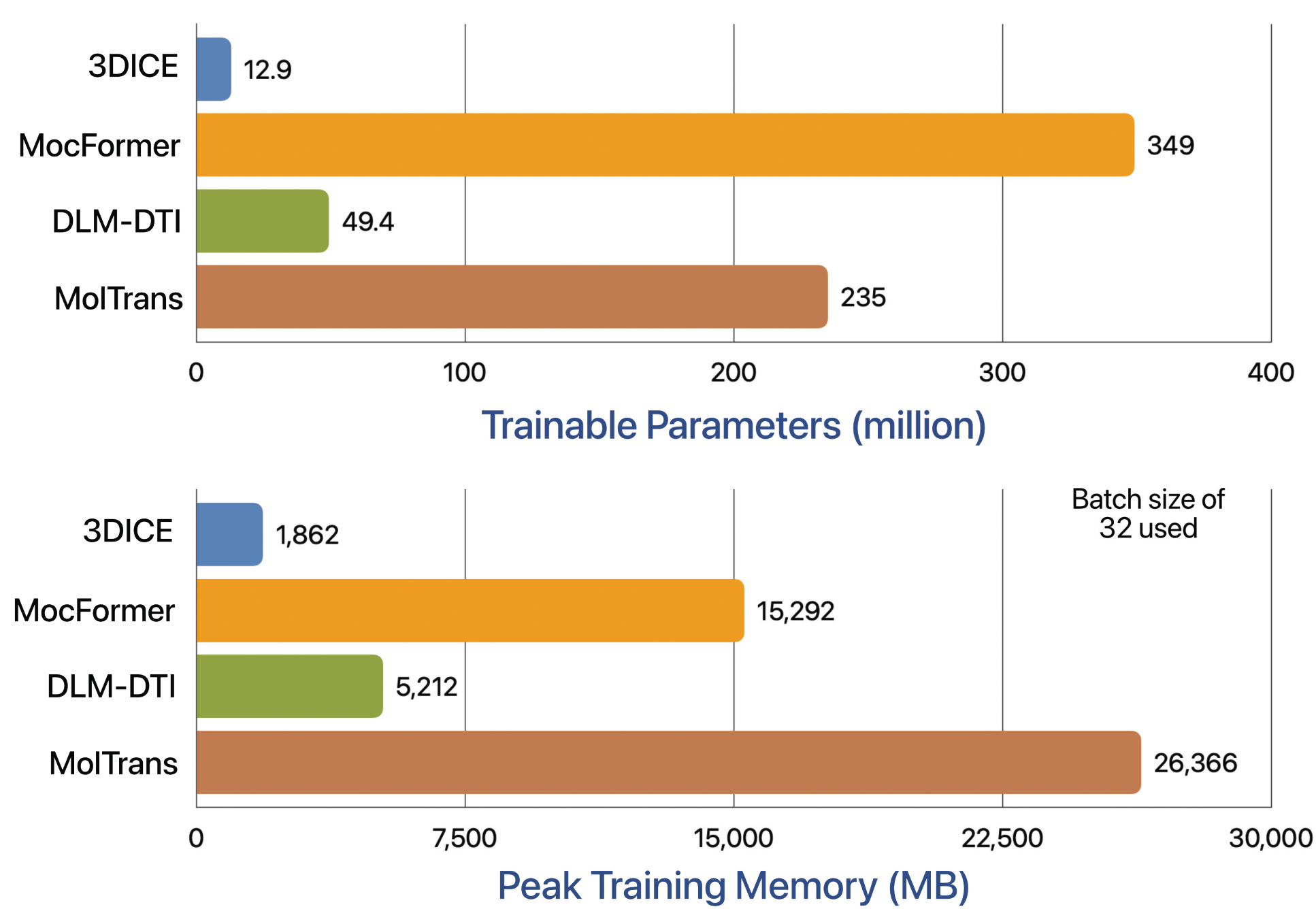


- Efficient learning of feature representations remains a **key bottleneck** in DTI prediction
- **Black box** predictions limit **trust** and **insight**
- Robust cold-start generalisation is **essential** for performance in real-world scenarios

Performance

- Comparable or superior performance to **state-of-the-art methods**
- Usage of 3D structural encoders and co-attention enables **effective generalisation** in real-world drug discovery scenarios

- Novel **faithfulness test** proves that attended areas are **critical** to model decision-making process
- Cross-attention weight maps can identify **putative interaction sites**
- Real-world usage simulated in a **case study** on the NSAID **celecoxib**



The diagram consists of four light blue rounded rectangular boxes arranged horizontally. Each box contains text describing a component of the model. The first box is titled 'Pre-training' in bold blue text and contains 'Transfers knowledge' and 'Better Generalisation'. The second box is titled '3D Geometry' in bold blue text and contains 'Captures binding-critical stereochemical cues'. The third box is titled 'Cross-Modal Fusion' in bold blue text and contains 'Explicitly models residue-atom interactions'. The fourth box is titled 'Interpretability' in bold blue text, with 'Not A Black Box' in bold purple text below it, and contains 'Reveals influential residues/atoms'.

Pre-training	3D Geometry	Cross-Modal Fusion	Interpretability
Transfers knowledge	Captures binding-critical stereochemical cues	Explicitly models residue-atom interactions	Reveals influential residues/atoms
Better Generalisation			Not A Black Box

Figure 3 displays the performance of MolTrans, DLM-DTI, and MocFormer on DrugBank, KIBA, and Cold-start datasets. The performance is evaluated using five metrics: ACC, PPV, TPR, F1, MCC, AUROC, and AUPRC. The results are presented in a table below the radar charts. The table compares the performance of the three models across the three datasets. The best results are bolded, and the second-best results are underlined. Five independent replications were performed (n = 5). Data is expressed as mean (std).

DrugBank

KIBA

Cold-start

ACC, PPV, TPR, F1, MCC, AUROC, AUPRC

Model	ACC	PPV	TPR	F1	MCC	AUROC	AUPRC
MolTrans	0.76205 (0.00372)	0.73491 (0.00552)	0.81959 (0.00025)	0.76359 (0.00703)	0.52764 (0.00689)	0.84060 (0.00267)	0.83459 (0.00387)
DLM-DTI	0.80135 (0.00636)	0.78714 (0.00811)	0.82643 (0.00782)	0.80627 (0.00587)	0.60350 (0.01256)	0.88026 (0.00464)	0.87446 (0.00529)
MocFormer	0.82610 (0.00847)	0.79685 (0.01246)	0.87493 (0.00494)	0.83400 (0.00644)	0.65487 (0.01538)	0.89803 (0.00468)	0.88385 (0.00700)
3DICE	0.84379 (0.00360)	0.85058 (0.00539)	0.83427 (0.00497)	0.84235 (0.00356)	0.68770 (0.00722)	0.91065 (0.00216)	0.90200 (0.00145)

The best results are bolded, while the second-best results are underlined. Five independent replications were performed (n = 5). Data is expressed as mean (std).

Faithfulness test • Deletion Curves

Protein deletion (positive samples)

Protein deletion (negative samples)

Drug deletion (positive samples)

Drug deletion (negative samples)

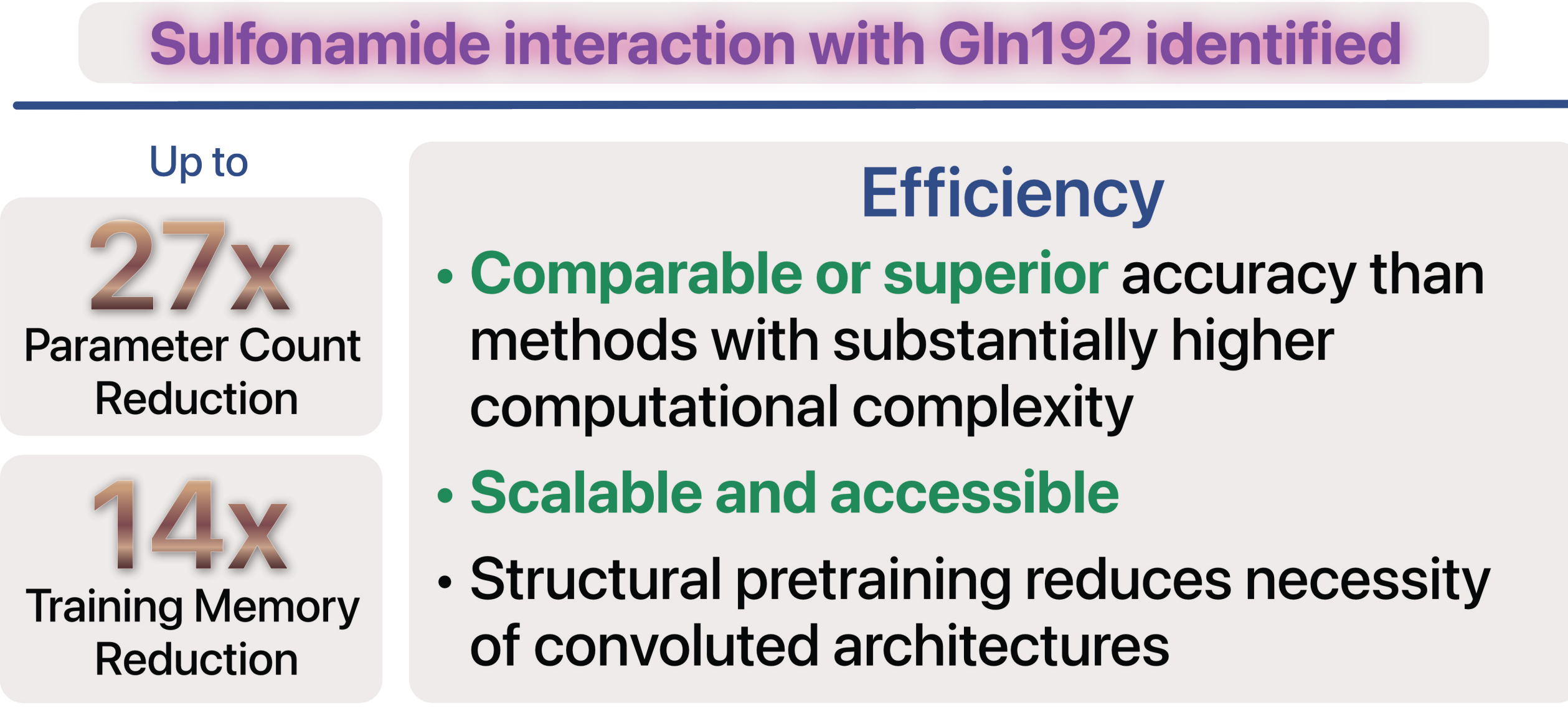
Fraction masked

Δmargin

selective

random

Model confidence decreases more when masking highlighted regions



The figure consists of three light pink rectangular cards arranged horizontally. Each card contains an icon, a title, a large numerical value, and a descriptive phrase.

- Cost:** Features a gold money bag icon with a dollar sign. The text reads "Cost", "US\$2.2B", "Per drug (avg)", and "Slow and resource-intensive".
- Scale:** Features a green test tube icon with a white plunger. The text reads "Scale", "35M+", "Purchasable compounds", and "Many-to-many interactions".
- Risk:** Features a yellow warning triangle icon with a black exclamation mark. The text reads "Risk", "10.8%", "Clinical success rate", and "Low probability of success".

The diagram illustrates the limitations of 1D molecular representations and Naïve Feature Fusion. It is divided into three main sections:

- Protein FASTA**: Shows a sequence of amino acids: ...TSPIMEELIT...HAAQVGLQDA...FHDHALMIF....
- Drug SMILES**: Shows a chemical structure: ...C(=NN1C1=CC=CS(N)....
- 1D Molecular Representations**: A central box with two red 'X' icons indicating limitations:
 - Encodes limited stereochemical information
 - Neglects important part of phenomenon
- Naïve Feature Fusion**: A box with a red 'X' icon indicating a limitation:
 - Simple Concatenation neglects modelling explicit cross-modal interactions

Uses a transformer to encode drug SMILES and protein sequences into embeddings

Hadamard-product interaction matrix + MLP

235 million trainable parameters

ProtBERT-distilled target encoder + ChemBERTa drug encoder
(both **1D sequence-based**)
Naïve concatenation + MLP
49.4 million trainable parameters

UniMol (drug 3D) + ESM-2 (1D protein sequence)
pre-trained encoders

Stacked self-attention + bilinear pooling + MLP

349 million trainable parameters

I am also prepared to present and explain...

Limitations and Future Work	More Data and Experiments
Pre-training logistics	Ablation studies
Alternative architectures	More compounds

```

graph TD
    A[Pre-trained 3D structural encoders] -- "+" --> B[Residue-atom co-attention]
    B -- "+" --> C[Novel Faithfulness Test]
  
```

Competitive on DrugBank / KIBA
+ cold-start scenario

Faithfulness tests prove significance
Highlighted regions verified as physical

Scalable drug–target interaction screening with interpretable binding cues

State-of-the-art performance with significant reduction in computational complexity

References available upon request.

All charts were created by the finalist.

Molecular graphics were generated using PyMOL.

AlphaFold, PDB, and RDKit logos were obtained from their official sources.