Standigm
AI for new drug R&D

QIP-ADMET™ (M4-ADMET)

A Modular Multi-Modal Approach to Drug Discovery

Executive Summary

The evolution of machine learning in drug discovery has crystallized around two main approaches: generalist and specialist models. While generalist models aim for broad applicability by integrating diverse data, their size can be computationally burdensome. Specialist models are efficient for specific tasks but may lack generalizability to novel chemical spaces. The **Standigm Foundation Model (M4)**, a **Modular Multi-Modal Model**, synthesizes the strengths of both, creating an interconnected network of models that share components and benefit from multitask learning across various data modalities. This approach results in more robust and generalizable representations. Key applications include **QIP-ADMET™(M4-ADMET) for comprehensive ADMET property prediction** and **M4-Affinity for efficient binding affinity prediction** in virtual screening. **Benchmarking on the TDC ADMET** dataset demonstrates the advanced performance of QIP-ADMET™, **surpassing the best models on the leaderboard** for several key ADMET properties, including a **~9% improvement in Lipophilicity prediction (MAE)** and **~7% improvement in PPBR and LD50 prediction (MAE)** over the leaderboard best (March 2025), and showing significant advancements over our previously published model. The framework also **supports post-training adaptation to user-provided data**, allowing for the incorporation of proprietary experimental results without requiring complete model retraining.

Introduction

The field of machine learning for drug discovery currently sees the development of two primary model archetypes: generalist and specialist models. Generalist models aim to address multiple, often heterogeneous, tasks within a single, large architecture. This approach is predicated on the hypothesis that integrating diverse, cross-domain knowledge can enhance generalization capabilities, particularly for novel domains and tasks. The recent success of Large Language Models (LLMs), which function as large-scale integrators of multiple modalities and tasks, has further catalyzed interest in this paradigm. However, the superiority of large generalist models over specialized counterparts is not universally guaranteed. Furthermore, their substantial size can impose significant computational

burdens, especially for large-scale inference tasks such as virtual screening. Certain tasks, like predicting basic physicochemical properties (e.g., solubility), may not necessitate the complexity of a generalist model, as lightweight specialist models often achieve comparable performance more efficiently.

Specialist models, conversely, are designed to solve a specific task or a small set of related tasks. Typically, multiple independent models, often traditional machine learning algorithms (e.g., random forests, gradient boosting trees) or Graph Neural Networks (GNNs), are trained on curated experimental datasets of relatively modest size. These models have proven valuable for specific property predictions. While generally lightweight and potentially high performing on their designated tasks (especially simpler ones like solubility or lipophilicity prediction), specialist models face limitations. Their reliance on smaller datasets and inherent lack of integrated cross-domain knowledge can restrict their generalizability to unseen chemical spaces or biological contexts. This limitation can be exacerbated by the relative simplicity of the underlying model architectures.

The Standigm Foundation Model (M4) Framework: Integrating Strengths

The Standigm Foundation Model (M4), which is short for a Modular Multi-Modal Model, is designed to synthesize the advantages of both generalist and specialist approaches. Instead of constructing a single, monolithic generalist model, the M4 framework generates an interconnected network of models that share a significant portion of their underlying components.

M4 architecture differs fundamentally from training specialist models in isolation. Models within the M4 framework interact and benefit from shared learning. For instance, a model predicting binding affinity can be enhanced by concurrently learning representations relevant to molecular property prediction, and vice versa. This is achieved through shared encoders and decoders for various data modalities (e.g., molecular structure, protein sequence). Individual task-specific models fuse embeddings from relevant modalities.

The models are trained across multiple tasks simultaneously. This strategy allows the shared modules (encoders, decoders) to benefit substantially from multitask learning and exposure to diverse, cross-domain data. Consequently, these modules are expected to develop more robust and generalizable representations compared to those learned by isolated specialist models. The M4 framework enables computational efficiency by allowing users to assemble only the necessary components for a given task. Simultaneously, it retains the advantages derived from cross-domain knowledge integration and multitask learning, fostering enhanced generalizability.

Figure 1 depicts the core design principle behind Standigm's M4 framework. Instead of training separate networks for every endpoint, M4 builds a library of shared components: (i) an embedder that decomposes each molecule into chemically meaningful fragments; (ii) a transformer encoder that learns universal atom- and fragment-level features; and (iii) a lightweight decoder that exposes these features to multiple downstream task heads. When a new property or activity prediction is needed, the relevant head simply plugs into the shared backbone, inheriting representations already optimized across many complementary tasks. This modularity cuts both training time and deployment cost while preserving the cross-domain generalization normally associated with much larger "all-in-one" models.

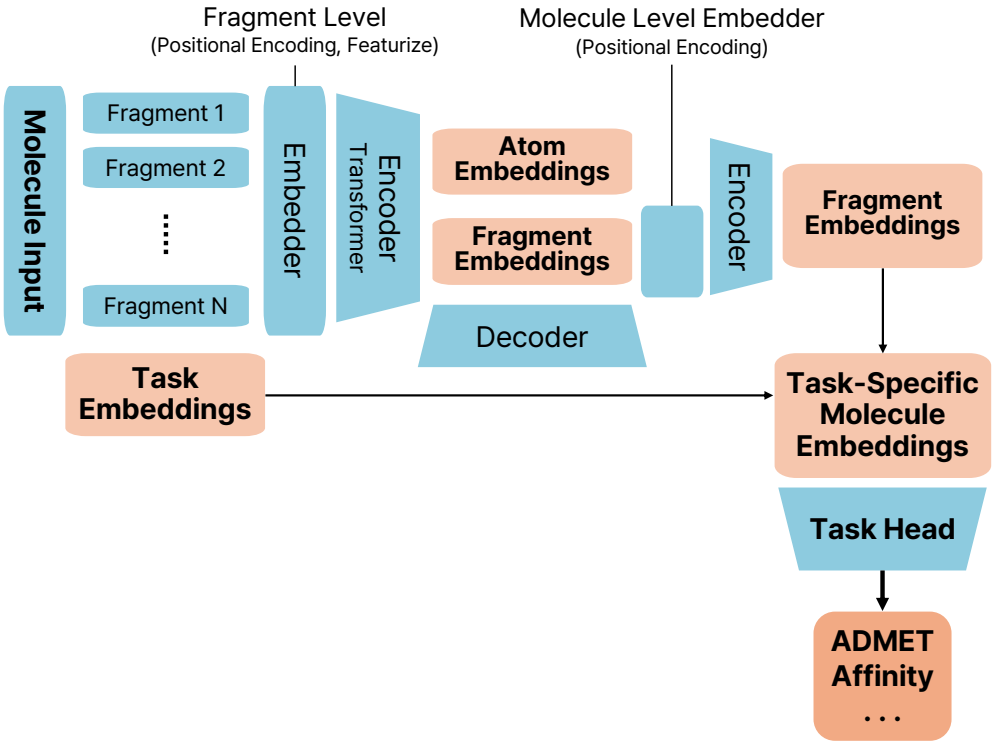

Figure 1. Modular Multi-Modal Model (M4) architecture with task-specific heads to generate predictions for diverse endpoints such as ADMET (QIP-ADMET™), binding affinity, etc.

Figure 2 summarizes the breadth of data used to pre-train M4. By mixing quantum-level metrics (e.g., dipole moments, ESP), biological read-outs (e.g., binding modes, ADMET panels), physical-chemistry constants (e.g., melting points, solubilities), and organic reaction outcomes, the model is exposed to chemical phenomena that span length-scales and disciplines. The resulting embeddings capture electronic, structural, and physicochemical regularities in a single latent space, giving subsequent specialist heads a strong inductive bias even when only modest task-specific datasets are available.

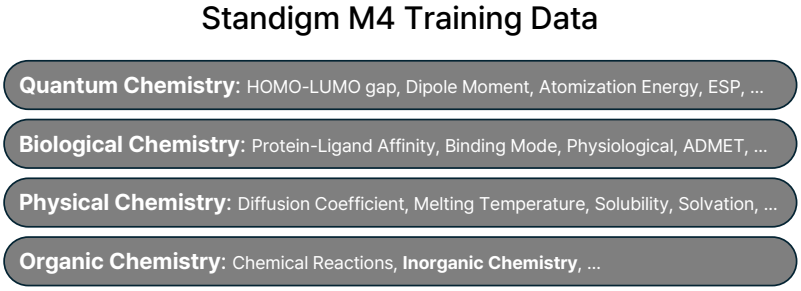

Figure 2. Cross-domain training corpus for the M4 foundation model.

Available Models within the M4 Framework

QIP-ADMET™ (M4-ADMET)

QIP-ADMET™ serves as the direct successor to our previous QIP-ADMET™ model. It leverages a novel molecular encoder embedded within the broader M4 framework, which also contributes to other predictive tasks such as binding affinity estimation. The underlying graph transformer architecture has been further refined for improved computational efficiency and predictive capability. A key enhancement is its ability to interpret molecular structures at the level of functional fragments, which is crucial for capturing substructure-property relationships relevant to ADMET profiles.

Training for QIP-ADMET™ was conducted on a comprehensive dataset comprising 43 distinct ADMET-related endpoints (Table 1). This dataset was meticulously compiled and curated from both public repositories and in-house data sources, ensuring broad coverage across Absorption, Distribution, Metabolism, Excretion, and Toxicity properties.

Table 1. ADMET properties supported by the QIP-ADMET™ model.

Absorption	Water solubility, PBS kinetic solubility, logP, logD at pH 7.4, Caco-2 permeability, Bioavailability, Colloidal self-aggregation
Distribution	Protein plasma binding rate (human), VDss (human), BBB, logBB, P-gp inhibition, P-gp substrate
Metabolism	CYP substrate and inhibition (1A2, 2B6, 2C19, 2C8, 2C9, 2D6, 3A4)
Excretion	Microsomal clearance (human, mouse, rat), Hepatic clearance (human, rat), Liver microsomal stability (human, mouse, rat), Plasma stability (human)
Toxicity	hERG inhibition, Acute toxicity LD50, Drug-induced liver injury, Ames mutagenicity, Carcinogenicity (mouse, rat)

Benchmarking on the TDC ADMET dataset highlights the advanced performance of the current QIP-ADMET™ model compared to both our previously published model [1] and the best models on the TDC leaderboard as of March 2025 (Table 2). Our continued development of the Standigm's M4 and the ADMET task has resulted in significant improvements. For instance, QIP-ADMET™ achieves superior performance in predicting Lipophilicity demonstrating a ~9% improvement. It also surpasses the leaderboard in predicting PPBR, a ~7% improvement, and LD50, another ~7% improvement. Furthermore, QIP-ADMET™ outperforms the leaderboard for Caco-2, HIA, BBB, and Clearance (Hepatocyte) predictions, indicating its robust and advanced capability across a range of crucial ADMET endpoints. These results underscore that QIP-ADMET™, powered by the latest iteration of the Standigm Foundation Model, M4, represents a significant advancement in ADMET prediction accuracy and reliability.

Table 2. QIP-ADMET™ performance on the TDC ADMET benchmark. Blue values surpass the current public leaderboard (Mar 2025), while bold values indicate improvements over the QIP-ADMET™ previously reported in 2024 [1].

Task		Metric	QIP-ADMET™		Best Leaderboard
			2024 [1]	2025	
Absorption	Caco-2	MAE ↓	0.268	**0.267**	0.276
	Bioavailability	AUROC ↑	0.725	0.713	0.753
	Lipophilicity	MAE ↓	0.442	**0.418**	0.460
	Solubility (aq)	MAE ↓	0.707	0.845	0.725
	HIA	AUROC ↑	0.995	0.993	0.990
	Pgp	AUROC ↑	0.927	**0.935**	0.938
Distribution	BBB	AUROC ↑	0.902	**0.926**	0.920
	PPBR	MAE ↓	7.364	**6.991**	7.505
	VDss	Spearman R ↑	0.608	0.688	0.713
Metabolism	CYP2C9 Inhibition	AUPRC ↑	0.787	**0.810**	0.859
	CYP2D6 Inhibition	AUPRC ↑	0.657	**0.672**	0.790
	CYP3A4 Inhibition	AUPRC ↑	0.871	0.864	0.916
	CYP2C9 Substrate	AUPRC ↑	0.519	0.376	0.441
	CYP2D6 Substrate	AUPRC ↑	0.665	**0.670**	0.736
	CYP3A4 Substrate	AUROC ↑	0.624	**0.635**	0.667
Excretion	Half Life	Spearman R ↑	0.529	**0.556**	0.576
	Clearance (Hepatocyte)	Spearman R ↑	0.508	**0.544**	0.536
	Clearance (Microsome)	Spearman R ↑	0.658	0.603	0.630
Toxicity	LD50	MAE ↓	0.562	**0.504**	0.541
	hERG	AUROC ↑	0.813	**0.863**	0.880
	Ames	AUROC ↑	0.857	0.827	0.871
	DILI	AUROC ↑	0.885	**0.904**	0.925

M4-Affinity

M4-Affinity is a new binding affinity prediction model developed upon the shared molecular and protein encoders within M4 framework. Its computational efficiency makes it particularly well-suited for deployment in large-scale virtual screening campaigns where throughput is a critical factor. In a lead-optimization scenario or when pipeline-specific modelling is required, M4-Affinity can be integrated within the Standigm BEST™ framework for binding affinity prediction.

Post-Training Adaptation

A common challenge arises when users possess proprietary experimental data generated from their specific pipelines and wish to incorporate this information into pre-trained models. Retraining large, complex models is often computationally prohibitive, whereas smaller, easily retrainable models may lack the required predictive power or generalizability. The Standigm Foundation Model, M4 framework addresses this by enabling adaptation to user-provided data without requiring complete model retraining. While benefiting from the powerful and generalizable representations learned by the core modules, the framework allows for post-training adaptation. When supplied with a relevant user dataset, the model's predictions are adjusted accordingly, and uncertainty estimates associated with these adapted predictions are also provided. For optimal performance and calibration of the adapted model, it is recommended that the distribution of the user-provided dataset closely resembles the chemical or biological space of the entities intended for prediction.

Conclusion

Standigm's M4 framework demonstrates that modularity and multi-task learning are not mutually exclusive. By sharing encoders and decoders across QIP-ADMET™(M4-ADMET), M4-Affinity and future task heads, the platform delivers (i) state-of-the-art accuracy on public benchmarks, (ii) computational efficiency that scales from desktop GPUs to cloud clusters, and (iii) the flexibility to incorporate proprietary data without full retraining.

Looking ahead, continued expansion of the cross-domain training corpus—especially with high-quality physicochemical and in-vivo datasets—will further enhance the model's extrapolative power.

Collectively, these advances position the M4 ecosystem as a foundational layer for AI-driven drug discovery, bridging early target identification (ASK™), compound design (BEST™),

and downstream developability assessment. Standigm welcomes collaborations to tailor M4 components to partners' unique data assets and therapeutic areas.

For more information and to explore how QIP-ADMET™ can enhance your drug discovery efforts, visit Standigm's QIP-ADMET™ page and Standigm's website.

Reference

[1] W. Chang, J. Kim, H. Ji and I. Joung, Quantum-Informed Molecular Representation Learning Enhancing ADMET Property Prediction, *J. Chem. Inf. Model.* 2024, 64, 13, 5028-5040.

Contact Us

For more information about Standigm and QIP-ADMET™ and how it can benefit you, please reach out to us:
- Email: business@standigm.com
- Website: www.standigm.com
- Latest QIP-ADMET™ White Paper: https://www.standigm.com/ai-saas/qip-admet

Legal Information

About Standigm

Standigm is a Korean AI drug discovery company with over 10 years of experience, backed by SK, Pavilion Capital, and Kakao Ventures, with $70M+ in funding and partnerships with global pharma. Standigm's proprietary AI platforms—Standigm ASK™ for target identification and Standigm BEST™ for compound design—power continuous generation of commercially valuable pipelines.