



Transfer Learning Without Chemical Amnesia for Predicting Natural Product Bioactivity Across Rare Scaffolds, New Targets, and Data-Sparse Therapeutic Domains

Rafael Mendoza^{1*}, Isabela Cruz¹, Carlos Lopez²

¹Department of Transfer Learning for Sparse Bioactivity, Faculty of Pharmacy, National Autonomous University of Mexico, Mexico City, Mexico.

²Department of Rare Scaffold and New Target Prediction, Faculty of Pharmacy, Monterrey Institute of Technology, Monterrey, Mexico.

ABSTRACT

Natural products remain a valuable but challenging source of bioactive compounds due to their structural diversity, stereochemical complexity, and uneven representation in public datasets, which often diverge from the distributions used to train contemporary machine-learning models. Transfer learning provides a promising solution for data-sparse prediction tasks, yet adaptation to new targets can inadvertently suppress information that is critical for chemically rare scaffolds, a phenomenon termed chemical amnesia that conventional accuracy metrics may fail to detect. This conceptual framework systematically examines the determinants of transfer reliability, including source-domain support, representation retention, biological-context conditioning, adaptation intensity, few-shot support composition, and domain-shift evaluation. The proposed Chemical-Memory Transfer Framework argues that successful transfer should be assessed not only by improvements in target-task performance but also by whether chemically meaningful information remains accessible throughout the adaptation process. The framework challenges the assumption that more extensive pretraining, broader multitask learning, or parameter-efficient fine-tuning is inherently beneficial, emphasizing instead that rare-scaffold prediction requires explicit evaluation of what knowledge is transferred, what is forgotten, and under which chemical and biological shifts. Prospective validation using independently sourced natural-product datasets and novel target contexts remains essential for establishing the practical utility of this approach.

Key Words: Natural products, Transfer learning, Molecular representation, Few-shot learning, Domain shift, Bioactivity prediction

eIJPPR 2025; 15(3):69-77

HOW TO CITE THIS ARTICLE: Mendoza R, Cruz I, Lopez C. Transfer Learning Without Chemical Amnesia for Predicting Natural Product Bioactivity Across Rare Scaffolds, New Targets, and Data-Sparse Therapeutic Domains. Int J Pharm Phytopharmacol Res. 2025;15(3):69-77. <https://doi.org/10.51847/h9vPTVjUoa>

INTRODUCTION

Natural products continue to contribute directly and indirectly to drug discovery, but their value is inseparable from the chemical and biological complexity that makes them difficult to model. Their structures occupy diverse, biosynthetically organized regions of chemical space that are not uniformly represented in conventional medicinal-chemistry datasets [1]. Contemporary discovery increasingly combines these molecules with computational

prioritization, yet predictive models inherit the boundaries of the chemical evidence on which they were trained. The practical question is therefore not simply whether transfer learning can improve prediction in low-data settings, but whether adaptation preserves information needed for chemistry that was uncommon in the source domain. This distinction shifts attention from transfer as parameter reuse toward transfer as selective preservation and reorganization of molecular information.

Corresponding author: Rafael Mendoza

Address: Affiliation: Department of Transfer Learning for Sparse Bioactivity, Faculty of Pharmacy, National Autonomous University of Mexico, Mexico City, Mexico.

E-mail: ✉ rafael.mendoza@unam.mx

Received: 26 January 2025; **Revised:** 29 March 2025; **Accepted:** 06 May 2025

This is an **open access** journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.



That concern begins at representation. Natural-product chemical space can look materially different depending on the molecular fingerprint used to describe it, so a scaffold may appear close, remote, redundant, or distinctive partly because of representational choice rather than chemistry alone [2]. This is especially consequential for transfer because source-domain support is usually estimated through the same representational machinery that later guides similarity, pretraining, task selection, or fine-tuning. A model that compresses biosynthetically meaningful distinctions before adaptation may never have a robust chemical memory to preserve, whereas a representation that separates rare structural motifs may expose sparsity that aggregate performance conceals. Rare chemistry can therefore be disadvantaged both by underrepresentation in the data and by under-resolution in the feature space.

Transfer remains attractive because molecular tasks often share information across related measurements, representations, and biological contexts. Recent studies show that downstream performance can improve when information is transferred across data-fidelity regimes, that molecular-property prediction depends on interacting choices of representation and model design, and that transfer can support drug-target interaction prediction under sparse conditions [3–5]. These findings justify transfer as a methodological opportunity, but not as an automatic solution to natural-product scarcity. This article therefore treats transfer as a controlled redistribution of learned molecular information and asks a stricter question: can a model adapt to a new target or therapeutic task without erasing chemically useful structure that may matter precisely because the downstream scaffold is uncommon? The proposed answer requires separating what is learned before transfer, what is retained during adaptation, and what remains reliable after chemical and biological domain shift.

Chemical amnesia in pretraining and fine-tuning

The term chemical amnesia is proposed here for a specific molecular-transfer failure: after adaptation, information previously useful for representing or predicting underrepresented chemistry becomes less accessible, less discriminative, or functionally unavailable. This construct is motivated by direct evidence that sequential molecular-property learning can exhibit catastrophic forgetting, with later learning degrading competence on earlier tasks [6]. Chemical amnesia is narrower than generic performance loss because it concerns loss of chemically relevant information across transfer stages. A decrease on a rare-scaffold test set is therefore suggestive but not diagnostic; domain mismatch, optimization instability, label noise,

insufficient source coverage, or an inadequate representation can produce a similar pattern.

Molecular pretraining makes the distinction important because it can create reusable representations before labeled data for the final task are available. Contrastive graph pretraining can yield transferable molecular representations, while knowledge-guided pretraining can incorporate structured chemical information into representation learning [7, 8]. Neither result establishes that all useful information survives later adaptation, and neither specifically validates retention for rare natural-product scaffolds. Pretraining should consequently be treated as an information-acquisition stage whose downstream value depends on what is learned, how broadly it covers chemical space, and how subsequent optimization modifies access to that information.

Transformer-based molecular prediction further illustrates why pretraining quality cannot be separated from adaptation design. Recent methodological synthesis shows substantial dependence on molecular encoding, pretraining objectives, dataset composition, architecture, and fine-tuning choices [9]. The proposed framework therefore separates representation acquisition from representation retention: useful source features may be learned yet lose downstream utility when adaptation is dominated by a narrow target task. Conversely, poor rare-scaffold performance need not indicate forgetting if the relevant chemistry was never learned. The distinction is operationally important because forgetting calls for retention-aware adaptation, whereas absent information calls for better source data or a more suitable representation.

Source-domain bias and scaffold imbalance

Natural-product transfer begins from an uneven source domain. Comparative chemoinformatics shows that known and readily obtainable natural products occupy characteristic regions of chemical space relative to commonly used synthetic or drug-like collections [10]. These differences do not imply that every natural product is out of distribution, but they make source composition consequential. If pretraining data are dominated by repeatedly sampled conventional chemotypes, rare natural-product scaffolds may receive weak local statistical support even when the total corpus is large. Dataset scale can therefore coexist with chemically important local scarcity.

Natural-product databases add a second layer of imbalance because observed frequency reflects curation and discovery history as well as chemistry. Classification of COCONUT structures shows that molecular features contain information associated with plant, fungal, or bacterial origin [11]. COCONUT itself aggregates

heterogeneous open natural-product resources, expanding accessible coverage while inheriting differences in annotation, provenance, and source completeness [12]. Such resources are therefore valuable for transfer research but should not be read as literal maps of chemical prevalence in nature. A scaffold uncommon in a database may be biologically uncommon, poorly sampled, inconsistently annotated, or historically neglected.

The problem is consequently better framed as source-domain bias plus scaffold imbalance rather than rarity alone. Analyses of plant chemical space reveal nonuniform taxonomic coverage, knowledge gaps, and unevenly characterized bioactive regions [13]. For transfer learning, these imbalances can influence which structures dominate source representations, which tasks appear related, and which examples enter low-data support sets. The proposed framework treats observed source frequency as an evidential property rather than a biological truth and separates structural novelty from annotation scarcity. This prevents failure on an uncommon scaffold from being prematurely attributed to chemical amnesia when inadequate exposure, sampling imbalance, or uncertain provenance may provide a more basic explanation.

Representation retention across transfer stages

If chemical memory is to be preserved, the information worth retaining must first be defined at a chemically meaningful level. Natural-product-oriented representation work shows that biosynthesis-informed substructures can be encoded explicitly in a lightweight, interpretable fingerprint [14]. This matters because a transfer procedure can preserve parameters without necessarily preserving distinctions that remain useful for rare chemistry. A retention-aware model should therefore be evaluated for continued access to task-relevant substructures and higher-order scaffold organization, rather than only for similarity of weights or latent embeddings. The relevant retained information will vary with the downstream endpoint and should not be imposed as a universal feature list.

Recent molecular pretraining methods further show that representation can extend beyond flat connectivity. Hierarchical self-supervised learning can organize information across structural levels, while self-conformation-aware pretraining can integrate conformational and substructural signals useful for selected downstream predictions [15, 16]. These results support the feasibility of chemically richer representations, not the stronger claim that such models resist forgetting during natural-product transfer. Richness may even increase the amount of potentially useful information that downstream adaptation can suppress if a small target dataset rewards only a narrow subset of the source representation.

Representation retention is therefore proposed as an independent transfer objective rather than a synonym for target-task accuracy. Catastrophic forgetting in molecular learning shows that successful acquisition of a later task does not guarantee preservation of earlier competence [6]. Conversely, interpretable natural-product representations show that chemically meaningful features can be made more directly interrogable than an undifferentiated latent vector [14]. A future retention test could compare source and adapted models on chemically stratified probes, rare-scaffold neighborhoods, or interpretable structural features while separately measuring target-task gain. Such tests would not prove mechanistic memory, but they would make the central trade-off observable: adaptation should improve task relevance without unnecessarily destroying access to chemistry that may matter under later domain shift.

Target-aware and task-aware adaptation

Rare chemistry is only one side of the transfer problem; the biological question also changes what molecular information should remain salient. Models that combine compound structure with textual bioassay descriptions show that assay context can contribute predictive information beyond structure alone [17]. For natural products, this supports adaptation that asks not only whether a molecule resembles source compounds, but also whether the new assay interrogates a comparable biological process and readout. Assay descriptions, however, can encode platform conventions, endpoint wording, or study-specific metadata. They should condition a model without being treated as mechanistic evidence.

Target awareness becomes more important when the destination protein is new or weakly represented. Protein-specific zero-shot drug-target prediction demonstrates that target information can support inference in settings without conventional target-specific training [18]. Such prediction does not establish physical binding, functional modulation, or therapeutic relevance. Within the proposed framework, target conditioning therefore narrows the inferential problem: it helps distinguish chemical novelty from biological novelty, allowing failure to be localized more precisely rather than attributed indiscriminately to an unfamiliar scaffold.

Task awareness also constrains how broadly information should be shared. Pharmaceutical multitask learning shows that related assay tasks can provide useful shared information, while practical benefit remains dependent on the task mixture and data context [19]. The manuscript therefore rejects an assumption that maximal task aggregation is necessarily protective. A broad multitask source may strengthen reusable features, but an unrelated

or dominant task can also redirect capacity away from distinctions needed downstream. **Table 1** converts these manuscript-derived assumptions into risks and remedies,

while **Figure 1** summarizes the resulting context-conditioned transfer logic.

Table 1. Assumption–risk–remedy matrix for context-aware transfer before low-data adaptation

Assumption	Transfer risk	Proposed remedy	Evidence status
Structure alone defines transferability	Rare scaffolds may be judged without the assay context that determines label meaning	Compare structure-only and assay-conditioned transfer before attributing failure to chemistry	Assay-context premise supported by [17]; remedy proposed
A new target is simply another sparse label set	Chemical and biological novelty become confounded	Condition explicitly on target information and report target novelty separately from scaffold novelty	Zero-shot target-conditioning premise supported by [18]; reporting rule proposed
More source tasks necessarily improve retention	Weakly related or dominant tasks may redirect shared capacity	Gate task sharing by relevance and inspect retained rare-chemistry competence after adaptation	Context-dependent multitask premise supported by [19]; retention test proposed
High downstream accuracy confirms successful transfer	Target gain can coexist with loss of chemically useful source information	Evaluate target-task gain together with representation-retention probes	Proposed synthesis grounded in preceding sections

Figure 1 is proposed as a context-conditioned map of what information should influence adaptation and what evidential status each influence has.

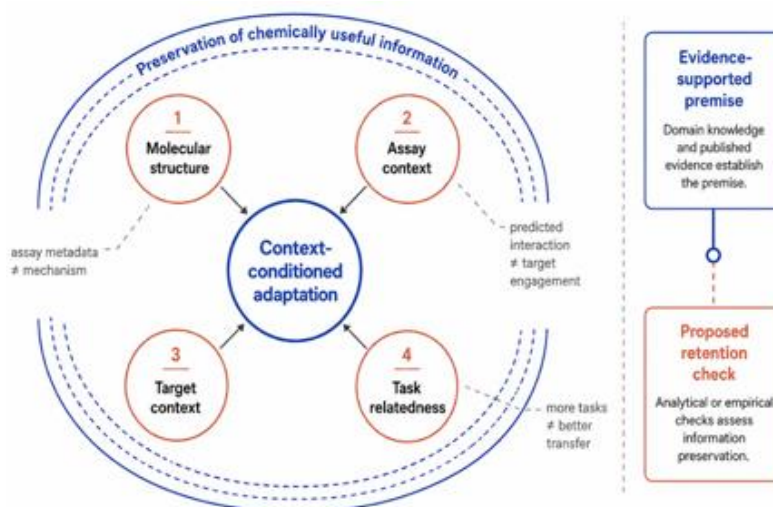


Figure 1. Context-conditioned transfer field for chemically rare molecules

Parameter-efficient and regularized transfer strategies

Once source and destination contexts have been characterized, adaptation intensity becomes a design variable rather than a default. A transferability map across molecular-property tasks shows that source choice materially affects whether transfer is beneficial, providing direct evidence that negative transfer cannot be excluded merely because two tasks concern molecules [20]. The proposed framework therefore places source-task screening before optimization. Similarity need not be represented by a single metric; chemical coverage, endpoint definition, target relation, and assay

comparability may each indicate different forms of transferability.

Parameter restriction and regularization are attractive because they can limit how far an adapted model moves from its pretrained state, but parameter efficiency should not be conflated with chemical-memory preservation. Multiview molecular learning combined with multitask optimization demonstrates that complementary representations and task objectives can change downstream predictive performance [21]. This supports preserving multiple information channels during adaptation, yet it does not show that freezing layers, adapters, penalties, or selective updates necessarily retain

rare-scaffold competence. Their value must be tested against an explicit retention endpoint.

Fine-tuning of pretrained protein and ligand language models further shows that adaptation can support prediction across unseen drugs, scaffolds, or targets in binding-affinity settings [22]. That result motivates controlled adaptation but does not establish functional activity or therapeutic effectiveness. For the present framework, a practical strategy is therefore comparative: evaluate full fine-tuning against restricted or regularized alternatives while holding the downstream task constant, then inspect both target-task performance and chemically stratified retention. A constrained model that preserves source features but cannot learn the destination biology has failed differently from an unconstrained model that adapts strongly while erasing useful chemical distinctions. Transfer should optimize that trade-off, not either objective alone.

Few-shot transfer to rare scaffolds and new targets

Few-shot learning addresses label scarcity, but label scarcity is not equivalent to chemical rarity. One-shot molecular learning established that useful bioactivity predictions can sometimes be obtained from very small support sets [23]. The important limitation for natural products is that a support set may be small yet chemically familiar, or larger yet poorly informative about the

destination scaffold. The proposed framework therefore treats support-set composition as part of the transfer mechanism: examples should be characterized by structural relation, target relation, assay comparability, and coverage of the rare chemistry being inferred.

Recent ligand-based few-shot methods reinforce the feasibility of learning under restricted activity data [24]. Their relevance does not imply that few-shot modeling solves simultaneous scaffold and target novelty. When the support compounds are structurally remote, the target is new, and assay conditions differ from the source, the model faces a compound shift rather than a single low-data problem. Under such conditions, apparently strong average performance on benchmark episodes may not estimate the uncertainty attached to the particular rare scaffold of interest.

In-context bioactivity prediction provides another route in which support examples condition predictions without conventional full retraining [25]. This can reduce adaptation of model parameters, but it does not remove dependence on what the context examples contain. **Table 2** therefore organizes few-shot transfer as a workflow of diagnostic checkpoints rather than as a single algorithmic choice. The central principle is that support data should reduce relevant uncertainty; merely increasing the number of examples is insufficient if they fail to bridge the chemical or biological gap.

Table 2. Workflow checkpoints for few-shot transfer under combined scaffold and target novelty

Workflow stage	Analytical question	Failure signal	Proposed action
Support-set diagnosis	Do examples cover the destination scaffold neighborhood and comparable assay meaning?	Small support set is chemically or biologically remote	Stratify support by scaffold, target, and assay relation; premise from [23]
Few-shot model selection	Can the method exploit sparse ligand evidence without assuming dense local coverage?	Performance depends on favorable support composition	Compare dedicated few-shot and simpler baselines; premise from [24]
Context construction	Do in-context examples contain information relevant to the destination task?	Added examples increase volume but not task relevance	Select support for informative relation rather than count alone; premise from [25]
Joint-novelty check	Are scaffold and target novelty being evaluated separately and together?	One familiar dimension masks failure on the other	Report single-shift and combined-shift performance; proposed
Retention check	Did adaptation or conditioning preserve useful rare-chemistry distinctions?	Target gain accompanies deterioration on retention probes	Compare target gain with chemically stratified retention; proposed

Evaluation under chemical and therapeutic domain shift

A model intended for rare scaffolds and new targets should be evaluated under the shifts it is expected to encounter. Real-world molecular out-of-distribution studies show that OOD behavior depends on how the domain shift is specified and that random IID evaluation is not a substitute for testing realistic novelty [26]. For natural-product transfer, chemical shift should therefore be operationalized explicitly—for example through scaffold, structural neighborhood, or source-collection separation—rather than inferred from a generic random split. Therapeutic-

domain shift should likewise be described through the destination evidence rather than a disease label alone: a new target family, assay platform, functional endpoint, or biological context can change what the prediction means. This prevents chemically familiar compounds tested in biologically unfamiliar settings from being misclassified as a purely chemical generalization problem.

Robustness studies further show that model performance and model ranking can change under OOD conditions [27]. This makes target-task accuracy on a familiar test distribution insufficient evidence for chemical-memory

transfer. Evaluation should distinguish at least two dimensions: distance from source chemistry and distance from source biological task. Assay or therapeutic-context differences can then be reported as modifiers of the biological axis rather than collapsed into a single notion of difficulty. No numeric threshold is proposed because the meaningful degree of shift depends on representation, endpoint, and intended use.

Large-scale drug-target benchmarking has also shown that methodological conclusions depend on evaluation construction [28]. The implication is not that one historic split should become a universal standard, but that

validation must prevent leakage of the relationships the model is supposed to extrapolate. **Figure 2** therefore expresses the manuscript-derived dual-shift logic as movable decision regions: stronger external evidence or successful retention tests may support more confident transfer within a region, whereas weak source coverage, assay mismatch, or uncertain retention should move the boundary toward guarded interpretation.

Figure 2 illustrates how chemical and biological novelty jointly determine the strength of evidence required before transferred predictions are interpreted.

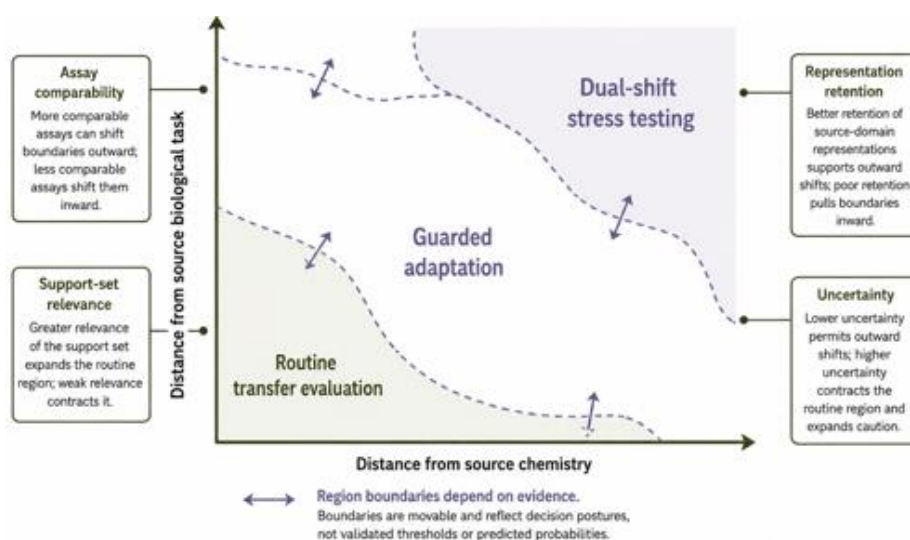


Figure 2. Dual-domain shift field for evaluating transferred bioactivity predictions The Chemical-Memory Transfer Framework

The proposed chemical-memory transfer framework integrates the preceding distinctions into a sequence of questions rather than a new predictive algorithm. First, the source domain is audited for chemical support: what parts of the destination scaffold, biosynthetic motif space, and representation are actually present before adaptation? Second, representation retention is evaluated independently of target-task gain. Third, the biological destination is characterized by target, assay, and task relation. Fourth, adaptation intensity is selected according to that joint evidence. Graph neural process work shows that generalization can deteriorate when target functions diverge and that additional adaptation can sometimes recover performance [29]. This supports treating adaptation need as conditional rather than fixed. A practical implementation would retain a transfer record containing the source chemical coverage, destination novelty, adaptation choice, and pre/post-adaptation retention checks. Such a record is proposed as an audit structure, not a validated score. Its purpose is to make competing explanations visible when a model improves on the destination task yet deteriorates on chemically stratified probes.

A second framework principle is that task relatedness acts as a gate on confidence in transfer, not as proof that transfer will succeed. Few-shot meta-learning studies show that task similarity can be crucial to structure-activity learning gains [30]. The framework therefore asks whether the source tasks contain biological and chemical relationships that are informative for the destination, while recognizing that any chosen similarity measure is itself incomplete. High apparent similarity cannot substitute for direct domain-shift testing, and low similarity may justify stronger adaptation or a decision not to transfer.

Third, chemical memory is not restricted to a single representation modality. Natural-product bioactivity classification from LC-MS/MS metabolomics demonstrates that analytically derived information can support prediction in settings where isolated molecular structures do not exhaust the available evidence [31]. The proposed framework consequently permits structure, target context, assay metadata, spectra, or other experimentally grounded modalities to contribute when they are genuinely available, but it keeps their evidential meanings separate. A spectral association is not a mechanism, and a predicted

structure-activity relation is not a measured pharmacological effect.

Finally, the framework requires prospective challenge in chemistry that is deliberately difficult rather than merely randomly held out. Targeted sampling of natural-product-like polyketide macrolide space shows that computational selection can be coupled to experimental screening in a constrained, unusual chemical domain [32]. This provides a model for validation philosophy, not proof of the present framework. A decisive test would compare adaptation strategies on independently sourced rare scaffolds and new targets, quantify target-task gain alongside retention loss, and experimentally evaluate selected predictions where feasible. The framework should be revised or rejected if retention-oriented constraints consistently fail to improve relevant generalization or prevent necessary adaptation.

Failure modes and applicability boundaries

Chemical-memory preservation cannot solve every molecular-prediction failure. Activity cliffs demonstrate that small structural changes can accompany large activity changes and that models with strong aggregate performance may still fail in these local regions [33]. Such error can occur even when the representation retains structural information, because the relevant structure-activity relationship is sharply discontinuous or poorly sampled. A retention-aware model should therefore report cliff-sensitive performance where pertinent rather than interpreting every rare-scaffold error as evidence of forgetting.

Bioactivity supervision can also be distorted before transfer begins. The construction of negative examples materially affects molecular bioactivity prediction, showing that label design can change what a model learns about inactivity [34]. Natural-product datasets may add further complications through heterogeneous assays or incomplete annotations. Consequently, deteriorating performance after transfer may reflect an altered label distribution rather than erased chemistry. Diagnosis should compare data provenance and endpoint definitions before assigning a memory-based explanation.

Finally, representation and train-test splitting themselves can substantially change apparent molecular-model performance [35]. The framework is therefore bounded to settings in which source and destination domains can be described with sufficient fidelity to distinguish representation failure, data shift, and adaptation-induced loss. It should not be used as a rhetorical label for any unsuccessful transfer experiment. Chemical amnesia is most defensible when useful source competence is demonstrated before adaptation, declines afterward on chemically relevant probes, and cannot be better explained by changed labels, incompatible measurements, unsupported chemistry, or ordinary stochastic variation.

Limitations

The framework inherits limitations from available natural-product data. Reviews of natural-product databases document substantial differences in coverage, curation, accessibility, and metadata [36]. A model cannot preserve information about chemistry that is absent or incorrectly represented in its source evidence, and observed database rarity should not be confused with biological rarity. The framework also leaves open how best to define chemically meaningful retention probes across different natural-product classes.

Uncertainty estimation can complement domain-shift testing but cannot guarantee reliability. Recent molecular-property work shows that uncertainty quantification remains challenging, particularly in difficult structure-activity settings [37]. Accordingly, high uncertainty may appropriately trigger caution, whereas low estimated uncertainty does not prove that an unfamiliar scaffold is safely within model competence.

The argument is also conceptual rather than prospectively validated. Evidence from specialized chemical domains such as targeted protein degraders illustrates that atypical molecular classes can require dedicated evaluation rather than automatic transfer of conventional modeling assumptions [38]. That analogy constrains generalization but is not direct evidence about natural products. Future tests should therefore span independent NP collections, multiple target families, heterogeneous assays, and prospective experimental confirmation. They should also examine whether conclusions change across molecular representations and definitions of scaffold novelty, because a retention signal that exists only under one encoding may reflect the probe rather than a stable property of the transferred model.

CONCLUSION

Transfer learning for natural-product bioactivity should be judged by more than whether a downstream metric improves. The proposed Chemical-Memory Transfer Framework reframes adaptation as a balance between acquiring task-specific information and preserving chemically meaningful source competence that may be needed for rare scaffolds, new targets, and later domain shifts. It separates source-domain support, representation retention, biological-context conditioning, adaptation intensity, few-shot support composition, and dual-domain evaluation so that distinct failure mechanisms are not collapsed into a single transfer score.

Chemical amnesia is proposed as a testable failure concept, not an established universal property of fine-tuning. Poor prediction may instead arise from missing source chemistry, assay mismatch, target dissimilarity, activity cliffs, label construction, representation choice, or

uncertainty. The framework therefore favors comparative adaptation experiments, chemically and biologically explicit holdouts, retention probes, and prospective challenge over claims based on random-split accuracy alone. Its value will depend on whether these distinctions improve diagnosis and validation in independent natural-product settings. If retention-oriented constraints do not improve relevant generalization, or if they impede necessary biological adaptation, the framework should be revised rather than protected from falsification.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

- [1] Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, et al. Natural products in drug discovery: Advances and opportunities. *Nat Rev Drug Discov.* 2021;20(3):200-16. doi:10.1038/s41573-020-00114-z
- [2] Boldini D, Ballabio D, Consonni V, Todeschini R, Grisoni F, Sieber SA. Effectiveness of molecular fingerprints for exploring the chemical space of natural products. *J Cheminform.* 2024;16(1):35. doi:10.1186/s13321-024-00830-3
- [3] Buterez D, Janet JP, Kiddle SJ, Oglic D, Lió P. Transfer learning with graph neural networks for improved molecular property prediction in the multi-fidelity setting. *Nat Commun.* 2024;15(1):1517. doi:10.1038/s41467-024-45566-8
- [4] Deng J, Yang Z, Wang H, Ojima I, Samaras D, Wang F. A systematic study of key elements underlying molecular property prediction. *Nat Commun.* 2023;14(1):6395. doi:10.1038/s41467-023-41948-6
- [5] Dalkiran A, Atakan A, Rifaioğlu AS, Martin MJ, Atalay RÇ, Acar AC, et al. Transfer learning for drug-target interaction prediction. *Bioinformatics.* 2023;39(Suppl 1). doi:10.1093/bioinformatics/btad234
- [6] Ranjan S, Singh SK. Overcoming catastrophic forgetting in molecular property prediction using continual learning of sequential episodes. *Expert Syst Appl.* 2025;267:125997. doi:10.1016/j.eswa.2024.125997
- [7] Wang Y, Wang J, Cao Z, Barati Farimani A. Molecular contrastive learning of representations via graph neural networks. *Nat Mach Intell.* 2022;4(3):279-87. doi:10.1038/s42256-022-00447-x
- [8] Li H, Zhang R, Min Y, Ma D, Zhao D, Zeng J. A knowledge-guided pre-training framework for improving molecular representation learning. *Nat Commun.* 2023;14(1):7568. doi:10.1038/s41467-023-43214-1
- [9] Sultan A, Sieg J, Mathea M, Volkamer A. Transformers for molecular property prediction: Lessons learned from the past five years. *J Chem Inf Model.* 2024;64(16):6259-80. doi:10.1021/acs.jcim.4c00747
- [10] Chen Y, Garcia de Lomana M, Friedrich NO, Kirchmair J. Characterization of the chemical space of known and readily obtainable natural products. *J Chem Inf Model.* 2018;58(8):1518-32. doi:10.1021/acs.jcim.8b00302
- [11] Capecchi A, Reymond JL. Classifying natural products from plants, fungi or bacteria using the COCONUT database and machine learning. *J Cheminform.* 2021;13(1):82. doi:10.1186/s13321-021-00559-3
- [12] Sorokina M, Merseburger P, Rajan K, Yirik MA, Steinbeck C. COCONUT online: Collection of Open Natural Products database. *J Cheminform.* 2021;13(1):2. doi:10.1186/s13321-020-00478-9
- [13] Domingo-Fernández D, Gadiya Y, Mubeen S, Healey D, Norman BH, Colluru V. Exploring the known chemical space of the plant kingdom: Insights into taxonomic patterns, knowledge gaps, and bioactive regions. *J Cheminform.* 2023;15(1):107. doi:10.1186/s13321-023-00778-w
- [14] Nollen LM, Meijer D, Sorokina M, van der Hooft JJJ. Biosynfoni: A biosynthesis-informed and interpretable lightweight molecular fingerprint. *J Cheminform.* 2025;17(1):136. doi:10.1186/s13321-025-01081-6
- [15] Zang X, Zhao X, Tang B. Hierarchical molecular graph self-supervised learning for property prediction. *Commun Chem.* 2023;6(1):34. doi:10.1038/s42004-023-00825-5
- [16] Qiao J, Jin J, Wang D, Teng S, Zhang J, Yang X, et al. A self-conformation-aware pre-training framework for molecular property prediction with substructure interpretability. *Nat Commun.* 2025;16(1):4382. doi:10.1038/s41467-025-59634-0
- [17] Schuh MG, Boldini D, Sieber SA. Synergizing chemical structures and bioassay descriptions for enhanced molecular property prediction in drug discovery. *J Chem Inf Model.* 2024;64(12):4640-50. doi:10.1021/acs.jcim.4c00765
- [18] Wang Y, Xia Y, Yan J, Yuan Y, Shen HB, Pan X. ZeroBind: A protein-specific zero-shot predictor with subgraph matching for drug-target interactions. *Nat Commun.* 2023;14(1):7861. doi:10.1038/s41467-023-43597-1

- [19] Ramsundar B, Liu B, Wu Z, Verras A, Tudor M, Sheridan RP, et al. Is multitask deep learning practical for pharma? *J Chem Inf Model*. 2017;57(8):2068-76. doi:10.1021/acs.jcim.7b00146
- [20] Yao S, Song J, Jia L, Cheng L, Zhong Z, Song M, et al. Fast and effective molecular property prediction with transferability map. *Commun Chem*. 2024;7(1):85. doi:10.1038/s42004-024-01169-4
- [21] Han X, Cai J, Bai C, Wu Z. Triview molecular representation learning combined with multitask optimization for enhanced molecular property prediction. *J Chem Inf Model*. 2025;65(10):5163-75. doi:10.1021/acs.jcim.5c00436
- [22] Gorantla R, Gema AP, Yang IX, Serrano-Morrás Á, Suutari B, Juárez-Jiménez J, et al. Learning binding affinities via fine-tuning of protein and ligand language models. *J Chem Inf Model*. 2025;65(22):12279-91. doi:10.1021/acs.jcim.5c02063
- [23] Altae-Tran H, Ramsundar B, Pappu AS, Pande V. Low data drug discovery with one-shot learning. *ACS Cent Sci*. 2017;3(4):283-93. doi:10.1021/acscentsci.6b00367
- [24] Eckmann P, Anderson J, Yu R, Gilson MK. Ligand-based compound activity prediction via few-shot learning. *J Chem Inf Model*. 2024;64(14):5492-9. doi:10.1021/acs.jcim.4c00485
- [25] Schimunek J, Luukkonen S, Klambauer G. MHNfs: Prompting in-context bioactivity predictions for low-data drug discovery. *J Chem Inf Model*. 2025;65(9):4243-50. doi:10.1021/acs.jcim.4c02373
- [26] Tossou P, Wognum C, Craig M, Mary H, Noutahi E. Real-world molecular out-of-distribution: Specification and investigation. *J Chem Inf Model*. 2024;64(3):697-711. doi:10.1021/acs.jcim.3c01774
- [27] Fooladi H, Vu TNL, Mathea M, Kirchmair J. Evaluating machine learning models for molecular property prediction: Performance and robustness on out-of-distribution data. *J Chem Inf Model*. 2025;65(19):9871-91. doi:10.1021/acs.jcim.5c00475
- [28] Mayr A, Klambauer G, Unterthiner T, Steijaert M, Wegner JK, Ceulemans H, et al. Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chem Sci*. 2018;9(24):5441-51. doi:10.1039/C8SC00148K
- [29] García-Ortegón M, Seal S, Rasmussen C, Bender A, Bacallado S. Graph neural processes for molecules: An evaluation on docking scores and strategies to improve generalization. *J Cheminform*. 2024;16(1):115. doi:10.1186/s13321-024-00904-2
- [30] Kötter A, Allenspach S, Grebner C, Matter H, Hiss JA, Schneider G, et al. Task-similarity is a crucial factor for few-shot meta-learning of structure-activity relationships. *Chembiochem*. 2024;25(19). doi:10.1002/cbic.202400095
- [31] Brittin NJ, Anderson JM, Braun DR, Rajski SR, Currie CR, Bugni TS. Machine learning-based bioactivity classification of natural products using LC-MS/MS metabolomics. *J Nat Prod*. 2025;88(2):361-72. doi:10.1021/acs.jnatprod.4c01123
- [32] Wilson DM, Driedger DJ, Liu DY, Keerthisinghe S, Hermann A, Bieniossek C, et al. Targeted sampling of natural product space to identify bioactive natural product-like polyketide macrolides. *Nat Commun*. 2024;15(1):2534. doi:10.1038/s41467-024-46721-x
- [33] van Tilborg D, Alenicheva A, Grisoni F. Exposing the limitations of molecular machine learning with activity cliffs. *J Chem Inf Model*. 2022;62(23):5938-51. doi:10.1021/acs.jcim.2c01073
- [34] Cáceres EL, Mew NC, Keiser MJ. Adding stochastic negative examples into machine learning improves molecular bioactivity prediction. *J Chem Inf Model*. 2020;60(12):5957-70. doi:10.1021/acs.jcim.0c00565
- [35] Yang K, Swanson K, Jin W, Coley C, Eiden P, Gao H, et al. Analyzing learned molecular representations for property prediction. *J Chem Inf Model*. 2019;59(8):3370-88. doi:10.1021/acs.jcim.9b00237
- [36] Sorokina M, Steinbeck C. Review on natural products databases: Where to find data in 2020. *J Cheminform*. 2020;12(1):20. doi:10.1186/s13321-020-00424-9
- [37] Kötter A, Singh K, Matter H, Hessler G, Grebner C. Upgrading reliability in molecular property prediction by robust quantification of uncertainty from machine learning models. *J Chem Inf Model*. 2025;65(20):10819-31. doi:10.1021/acs.jcim.5c00464
- [38] Peteani G, Huynh MTD, Gerebtzoff G, Rodríguez-Pérez R. Application of machine learning models for property prediction to targeted protein degraders. *Nat Commun*. 2024;15(1):5764. doi:10.1038/s41467-024-49979-3