Journal of Liquid Chromatography and Gas Chromatography, Europe
https://technicalreport.org/lcgceu/

LC GC Eu., Vol. 38, No.2, 17-26, 2025

Technical Reports Press

RESEARCH ARTICLE

Reliability-Gated Similarity Cartography for Transferable Quantitative Structure–Retention Modelling in Chromatographic Analytical Science

Mark A. Schurko[1*]

[1]*Department of Chemistry, Northwestern University, Evanston, IL, USA.*
[*]*Corresponding author*

Corresponding author:
maschurko@northwestern.edu

Abstract

Retention prediction is increasingly used as an orthogonal decision layer in LC–MS, GC–MS, metabolomics, environmental screening, pharmaceutical analysis, and forensic toxicology, yet retention values remain conditional on chromatographic mode, stationary phase, mobile phase, ionization state, derivatization status, and chemical-class coverage. This paper asks whether heterogeneous retention repositories can be organized into transferable quantitative structure–retention relationship (QSRR) evidence without permitting unsupported predictions to influence compound identification. Reliability-gated similarity cartography (RG-SC) answers this question by arranging RPLC, HILIC, GC, mixed-transfer, chiral, condition-diverse RPLC, drug-discovery screening, and forensic HRMS retention records as mode-resolved analytical evidence. Local neighbourhoods are accepted only when fingerprint similarity, distribution-coefficient proximity, retention-factor proximity, chemical-nature agreement, dual structural/retention screening, and second-dominant-interaction compatibility point to the same chromatographic environment. Retention estimates pass forward only when the local chemical map, chromatographic-system context, validation behaviour, and neighbourhood density are mutually consistent. RG-SC changes QSRR from unrestricted global prediction into a defensible analytical decision process: it separates incompatible systems, identifies sparse or stereochemically sensitive domains, supports retention-assisted candidate filtration, and withholds predictions when evidence is insufficient. Transferable retention modelling is therefore most reliable when handled as evidence-gated local cartography rather than extrapolation from pooled retention values.

Keywords: analytical chemistry; chromatographic retention; quantitative structure–retention relationship; retention-time prediction; LC–MS; GC–MS; chemometrics; instrumentation; compound identification; applicability domain.

Received: 23 November 2025
Accepted: 23 November 2025
Available online: 23 November 2025

1. Introduction

Chromatographic retention is a compact analytical signal that records how a molecule interacts with the stationary phase, the mobile phase, and the operating conditions of a separation. In liquid chromatography, retention responds to hydrophobicity, polarity, ionization state, hydrogen bonding, steric exposure, and stationary-phase selectivity. In gas chromatography, retention indices and retention times are shaped by volatility, temperature program, stationary-phase interactions, and derivatization status. Because these properties are partly independent of accurate mass and fragmentation behaviour, retention is valuable as an orthogonal constraint during LC–MS and GC–MS identification. A candidate that fits elemental formula and fragment evidence may still be chromatographically implausible under the measured system, while a candidate with moderate spectral ambiguity may become more credible when its retention behaviour matches the analytical context.

Quantitative structure–retention relationship modelling emerged from the broader chemometric idea that molecular descriptors encode interactions responsible for chromatographic migration. Classical QSRR studies related retention to lipophilicity, molecular size, polarizabil-

Cite as: M. A. Schurko (2025). Reliability-Gated Similarity Cartography for Transferable Quantitative Structure–Retention Modelling in Chromatographic Analytical Science. LC GC Eu., 38(2) (2025) 17-26.

ity, electronic distribution, acidity or basicity, and topological structure [1–3]. This foundation remains important because retention models are not only prediction tools; they also express a mechanistic interpretation of how molecular structure interacts with a separation system. For analytical chemistry and instrumentation, a useful QSRR model must therefore connect molecular representation with chromatographic meaning rather than produce a numerical retention estimate without explanation.

The scientific challenge is that retention is not a universal molecular constant. A compound measured by reversed-phase liquid chromatography (RPLC) can be controlled mainly by hydrophobic and dispersive interactions, whereas the same compound under hydrophilic interaction liquid chromatography (HILIC) can be controlled by polar partitioning, ionic effects, and hydrogen bonding. GC retention-index data are governed by a different scale of evidence, and the generalized retention-index concept remains central to the transfer of GC retention information across temperature-programmed separations [4, 5]. Even within RPLC, column chemistry, gradient slope, dwell volume, organic modifier, pH, buffer identity, and temperature can shift retention order. A transferable retention model must therefore distinguish genuine chemical similarity from accidental numerical proximity.

Large public and curated retention repositories have made QSRR more powerful, but they have also made domain control more urgent. METLIN provides broad reversed-phase retention-time coverage for machine-learning prediction [6]. Retip demonstrates how retention prediction can assist compound annotation in both HILIC and RPLC analytical procedures [7]. PredRet shows that direct mapping between chromatographic systems can make laboratory-specific retention information more reusable [8]. FiehnLib and Golm collections strengthen GC–MS metabolomics by combining spectral and retention-index evidence [9]. HILIC-specific modelling confirms that polar retention prediction requires attention to chemical class and instrumental conditions [10]. These repositories collectively show that retention prediction is now a cross-platform analytical problem rather than a narrow single-column modelling task.

Modern molecular representation has widened the modelling options. ECFP circular fingerprints provide efficient substructure similarity and remain useful for local-neighbour selection [11]. Partial least squares, random forests, gradient boosting, support-vector regression, graph neural networks, and transformer-based models can all learn structure–retention relationships when chromatographic coverage is adequate [12–18]. Recent calibration and deep-learning studies indicate that retention-time prediction can improve across LC methods when system-specific effects are corrected, but such correction does not automatically solve chemical-domain mismatch [19, 20]. High-capacity models still need applicability-domain evidence, uncertainty control, and chromatographic interpretation.

Retention prediction also intersects with the broader problem of confidence in small-molecule identification. Reporting standards and identification-confidence schemes emphasize that analytical assignments should communicate the strength and limits of supporting evidence [21–23]. Spectral tools such as MetFrag and SIRIUS improve structure elucidation from high-resolution and tandem mass spectra, yet candidate lists may remain ambiguous when spectral evidence is incomplete or compounds are isomeric [24, 25]. Retention evidence can reduce false positives, but only when its own reliability is clear. A retention estimate that is outside the model domain can be more harmful than no estimate because it may eliminate the correct candidate or promote a chemically plausible but chromatographically unsupported structure.

This study addresses a specific research question: can heterogeneous chromatographic retention records be organized into local, mode-aware similarity maps that support transferable QSRR decisions while explicitly withholding predictions outside the reliable domain? The question differs from ordinary retention-time prediction because the objective is not to maximize a single global accuracy statistic. The objective is to decide when retention evidence is strong enough to influence analytical identification and when it should remain advisory or be excluded.

Reliability-gated similarity cartography answers this question by treating each query compound as a local analytical problem. Neighbours are selected only when structural, physicochemical, chromatographic, chemical-class, and secondary-interaction evidence support the same local environment. Predictions then pass through a reliability gate that evaluates validation behaviour, neighbourhood density, transfer compatibility, and applicability-domain position. Each decision is described in chromatographic terms so that analytical chemists, separation scientists, and instrumentation researchers can inspect why a retention estimate is accepted, downgraded, or withheld during retention-assisted compound identification.

2. Materials and analytical design

2.1. Scope and analytical positioning

The study belongs to Chemistry, Analytical, Spectroscopy/ Instrumentation/ Analytical Sciences because retention is treated as an instrumental analytical signal that supports compound identification, method transfer, and decision control. The platform scope includes RPLC, HILIC, GC, mixed HILIC/RPLC transfer systems, chiral retention settings, isocratic records, and condition-diverse reversed-phase records. The retention corpus is not a single experimental run or a pooled numerical table; it is a mode-resolved chromatographic evidence system in which each record class is assigned to the analytical role for which it can support prediction.

The analytical aim is to protect the interpretability of retention evidence. A compound can be inside the molecular descriptor range of a global model while still being outside the chromatographic domain of the target method. RG-SC therefore defines reliability through four conditions: compatible chromatographic system, coherent chemical neighbourhood, adequate validation behaviour, and sufficient local density. These conditions distinguish the procedure from a conventional QSRR model that returns a prediction for every query regardless of support.

2.2. Mode-resolved retention corpus

The mode-resolved retention corpus separates evidence into record classes before any local modelling is performed. RPLC, HILIC, GC, mixed transfer, chiral, isocratic, and condition-diverse records are retained as distinct analytical environments. This separation prevents a model from treating retention time, retention factor, retention index, and transfer-calibrated retention as if they were the same response variable.

The retention corpus in Table 1 defines the paper's evidence base. Its main implication is that retention records are not interchangeable simply because they contain a molecular structure and a retention value. GC records remain retention-index evidence, HILIC records remain polarity-sensitive retention evidence, and RPLC records remain tied to column chemistry and transfer status. This organization gives the later reliability gate a meaningful input: it can reject a prediction because the nearest records come from the wrong analytical environment rather than because the structure itself is unfamiliar.

The retention-corpus map in Figure 1 reinforces the same design principle. The RPLC block is intentionally larger because reversed-phase records are more numerous and more heterogeneous, whereas HILIC and GC blocks are separated because their retention mechanisms differ from RPLC. The mixed/transfer block functions as a bridge rather than as a universal correction layer. This distinction matters in practice: a transfer record can support prediction only after the local chemistry and the target mode agree.

Table 1: Curated retention evidence layers.

Database or record class	Mode	Role in RG-SC
METLIN retention records	RPLC	Broad reversed-phase learning and small-molecule annotation support.
NIST-type retention-index records	GC	Retention-index evidence for volatile or derivatized compounds.
Golm metabolome records	GC	GC–MS metabolomics retention-index and spectral context.
Retip records	HILIC/RPLC	Practical retention prediction for metabolomics annotation.
Pathogen Box records	HILIC/RPLC	Drug-discovery and bioactive-compound retention diversity.
Wilson-type RPLC records	RPLC	Column-selectivity and hydrophobic-subtraction evidence.
Tan-type RPLC records	RPLC	Reversed-phase transfer and descriptor comparison.
PredRet-type records	HILIC/RPLC	Direct mapping between chromatographic systems.
Fiehn HILIC records	HILIC	Polar-metabolite retention and HILIC-specific learning.
Hall-type records	RPLC	Supplementary reversed-phase retention coverage.
RIKEN records	RPLC/HILIC	Metabolomics-oriented retention and annotation support.
Isocratic retention records	RPLC	Retention-factor interpretation under controlled composition.
Chiral retention records	Chiral LC	Stereochemical applicability-domain testing.
Van Laethem-type records	RPLC	Condition-diverse reversed-phase transfer evidence.
Aicheler-type RPLC records	RPLC	Instrument-specific reversed-phase retention modelling.
Forensic HRMS screening records	HILIC/RPLC	Toxicology-oriented retention-assisted filtering without unrestricted extrapolation.

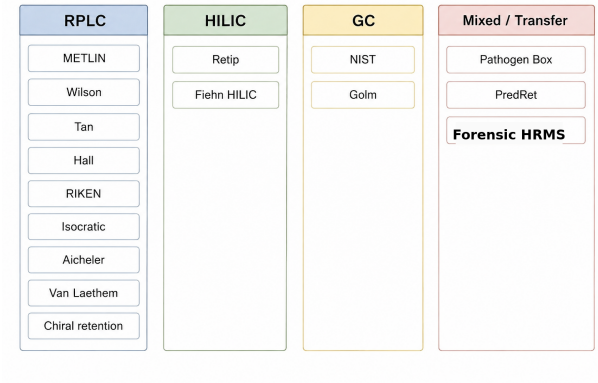


Figure 1: Mode-resolved retention repositories.

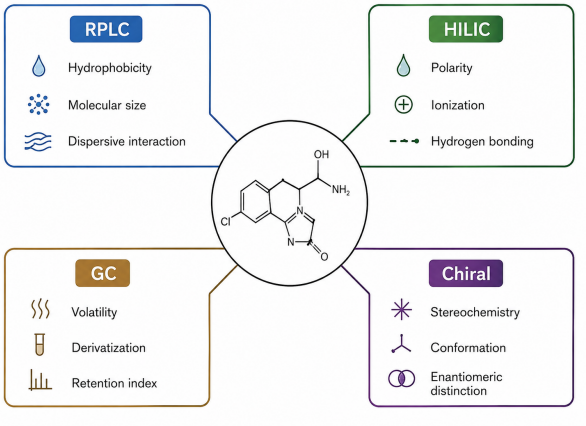


Figure 2: Descriptor emphasis by mode.

2.3. Molecular representation

Molecular representation is constructed in a mode-aware manner. Fingerprints supply substructure similarity, while physicochemical descriptors supply retention-relevant interpretation. RPLC maps emphasize hydrophobicity, molecular size, polar surface contribution, ionization state, and column-selectivity terms. HILIC maps emphasize polarity, ionization, hydrogen-bonding capacity, and zwitterionic behaviour. GC maps emphasize volatility-related structure, derivatization status, and retention-index compatibility. Chiral maps require stereochemical descriptors because enantiomers and diastereomers may be indistinguishable under ordinary fingerprints while remaining chromatographically distinct.

The descriptor-centre display in Figure 2 explains why one descriptor set cannot be assumed to serve all chromatographic modes equally. Hydrophobicity is informative for RPLC, but it cannot replace polarity and ionization terms in HILIC or volatility-related terms in GC. The practical result is a more conservative modelling sequence: similarity is judged relative to the retention mechanism being used, not relative to a universal chemical fingerprint alone.

2.4. Similarity-evidence layer

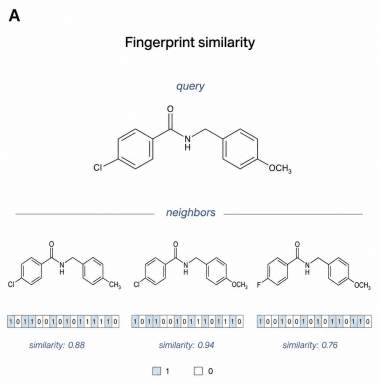
RG-SC defines a neighbour as a compound supported by multiple forms of similarity. Structural fingerprints are necessary but not sufficient. Distribution-coefficient proximity, retention-factor proximity, chemical-nature agreement, dual filtering, and second-dominant-interaction compatibility are added to reduce false neighbourhoods. This design is especially important when a small substituent changes ionization behaviour or when two scaffolds differ but interact with the chromatographic system through the same dominant retention mechanism.

The evidence streams in Table 2 create a practical safeguard against chemically plausible but chromatographically weak neighbours. Fingerprint similarity answers whether compounds look structurally related, whereas distribution and retention proximity ask whether they are likely to behave similarly under the operating conditions. Chemical-nature agreement prevents acid/base or neutral/zwitterionic mixing from being treated as harmless. Second-dominant-interaction compatibility gives the method a route to capture column selectivity, hydrogen bonding, and shape effects that are often decisive in real separations.

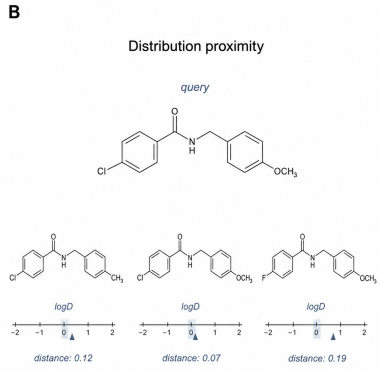
The six-panel evidence structure in Figures 3 and 4 show why the method cannot be reduced to fingerprint matching. A query may share a scaffold with a neighbour but differ in ionization and retention factor; another query may share retention behaviour through polarity or hydrogen bonding despite lower scaffold similarity. RG-SC uses these streams as cumulative evidence, so a local map becomes credible only when structural, physicochemical, and chromatographic signals point in the same direction.

Table 2: Similarity evidence streams.

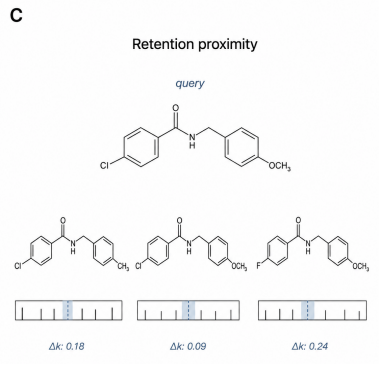
Similarity stream	Operational definition	Retention relevance
Fingerprint similarity	Substructure-level comparison based on molecular fingerprints.	Captures scaffold, substituent, and local structural relatedness.
Distribution-coefficient proximity	Comparison of logD behaviour under chemically relevant conditions.	Captures ionization-aware lipophilicity and pH-sensitive partitioning.
Retention-factor proximity	Comparison of compatible retention-factor behaviour when records are available.	Links neighbourhood selection to observed chromatographic behaviour.
Chemical nature agreement	Class agreement among acidic, basic, neutral, zwitterionic, and amphiphilic analytes.	Reduces misleading local maps across different ionization regimes.
Dual filtering	Sequential structural and retention-related filtering before local-map construction.	Requires both chemical plausibility and chromatographic compatibility.
Second-dominant-interaction compatibility	Similarity based on hydrophobic, steric, hydrogen-bonding, dipolar, and selectivity coefficients.	Preserves secondary interactions between solutes, mobile phase, and stationary phase.



(a) Fingerprint.

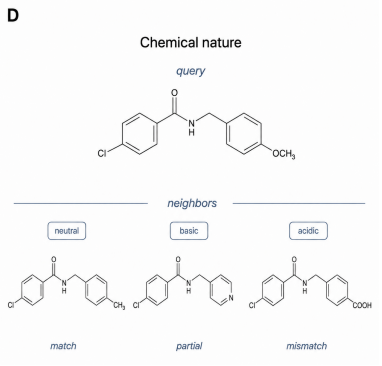


(b) Distribution.

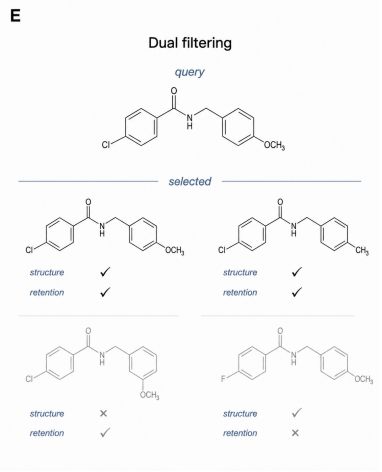


(c) Retention.

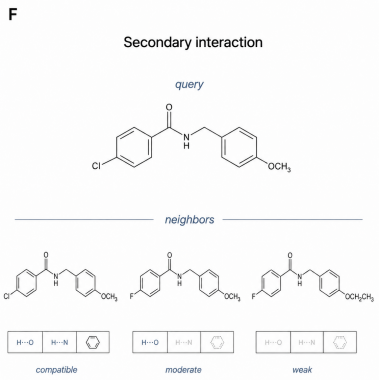
Figure 3: Similarity streams for local mapping.



(a) Chemical class.



(b) Dual filter.



(c) Secondary interaction.

Figure 4: Similarity streams for local mapping.

2.5. Local retention maps and reliability gate

For each query compound, RG-SC builds a local retention map inside the relevant mode-resolved record class. The map may be RPLC-focused, HILIC-focused, GC-focused, mixed transfer-oriented, chiral, or condition-diverse. Candidate neighbours are screened through the similarity-evidence layer before the prediction model is allowed to operate. When the neighbourhood is dense and chemically coherent, a local predictor can be used. When the neighbourhood is sparse, chemically inconsistent, or drawn from incompatible chromatographic systems, the gate reduces the status of the prediction or withholds it.

The reliability gate is a decision-control component. It evaluates trend agreement, average error, large-error sensitivity, cross-validation stability, relative error, external predictive ability, robust central error, neighbourhood density, and transfer consistency. This design follows validation principles in QSAR and QSRR while translating them into chromatographic terms [26, 27]. The gate is intentionally conservative because retention predictions are often used to support real analytical

decisions.

The validation layer in Table 3 prevents a single attractive statistic from dominating interpretation. High trend agreement is useful, but it does not guarantee acceptably small errors near identification boundaries. Average error is informative, but it can hide rare large failures. Relative error is needed when short and long gradients are compared. External predictive ability is the final check because the analytical use case involves compounds not used to construct the local map.

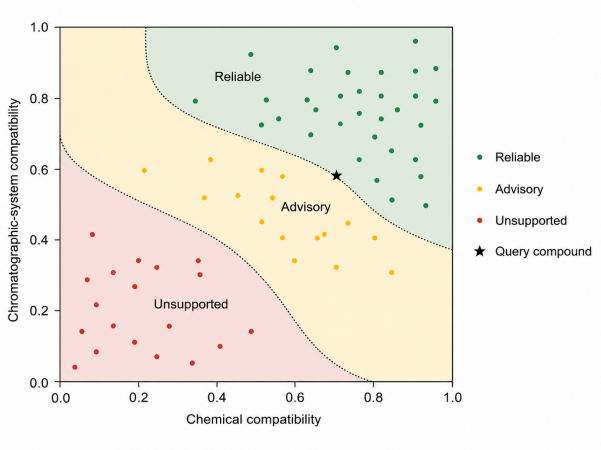


Figure 5: Reliability zones for retention prediction.

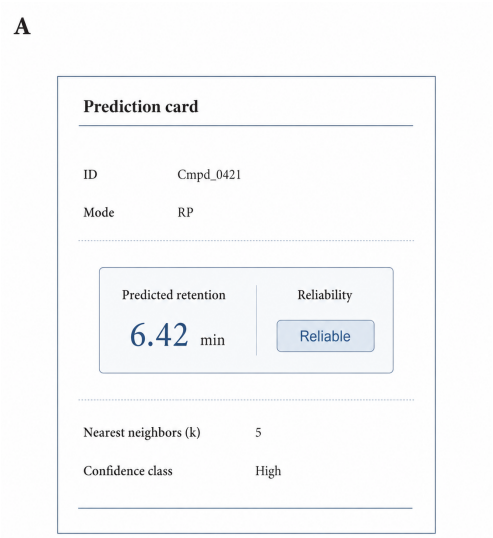
The reliability landscape in Figure 5 translates validation and domain evidence into three practical zones. The accepted zone represents dense and compatible local evidence. The advisory zone represents partial evidence that can inform interpretation but should not determine an identity alone. The unsupported zone represents extrapolation outside the reliable chemical or chromatographic domain. This triage is essential because an unsupported numerical output can appear more precise than it truly is.

The instrument-facing report in Figures 6 and 7 show how the procedure can be used by analysts at the point of interpretation. The predicted value is accompanied by a system label, local-neighbour count, validation state, applicability-domain position, and final gate status. This presentation supports laboratory review because the analyst can see why a prediction is accepted, downgraded, or withheld before using it for candidate prioritization.

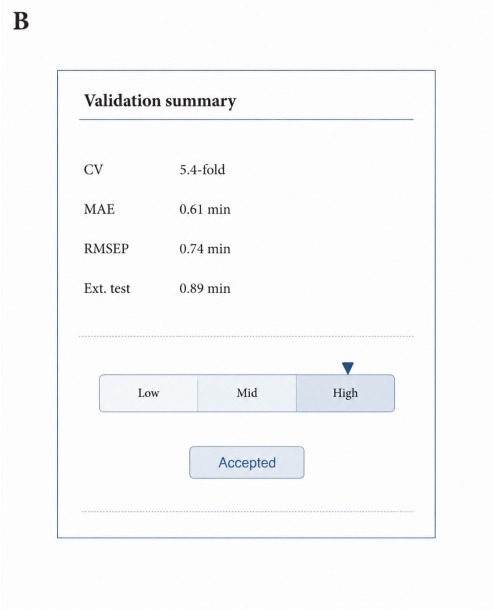
2.6. Analytical assessment protocol

The assessment protocol uses three tasks. The first task examines cross-system transfer between RPLC, HILIC, GC, and mixed HILIC/RPLC records. METLIN, Wilson-type, Tan-type, Hall-type, RIKEN, isocratic, Aicheler-type, and Van Laethem-type records support reversed-phase evaluation; Retip, Fiehn HILIC, Pathogen Box, and PredRet-type records support HILIC or mixed-mode evaluation; NIST-type and Golm records support GC retention-index evaluation. The second task evaluates sparse-class local prediction for underrepresented, stereochemically sensitive, or instrumentally difficult compounds. The third task evaluates retention-assisted candidate filtration by testing whether retention evidence can reduce candidate lists only when the reliability gate permits its use.

This protocol creates an assessment structure that is different from a single pooled accuracy value. RG-SC is judged by whether it preserves system identity, improves neighbour quality, identifies unsupported domains, and produces an interpretable decision state for candidate filtration. These are the central outcomes needed for analytical deployment.



(a) Prediction card.



(b) Validation summary.

Figure 6: Instrument-facing reliability report.

3. Results and discussion

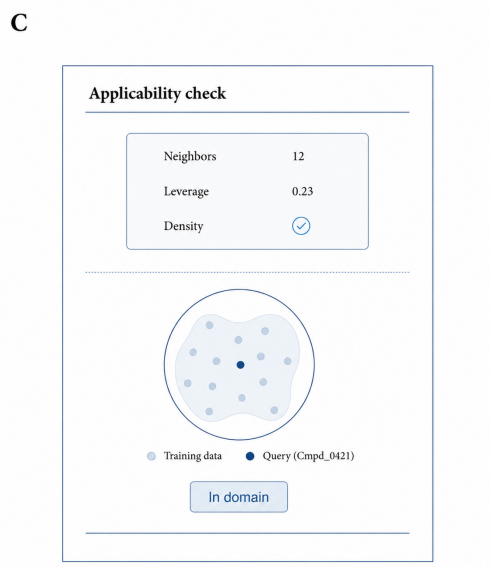
3.1. Organization of chromatographic information

The mode-resolved corpus changes the modelling problem from pooled regression to chromatographic evidence management. METLIN-like reversed-phase records provide broad small-molecule coverage, but they cannot automatically support HILIC or GC prediction. NIST-type and Golm records supply retention-index evidence that has a different scale and mechanistic basis from LC retention time. Retip, Fiehn HILIC, and PredRet-type records supply polarity-sensitive and transfer-oriented evidence. Chiral and condition-diverse records test whether the local map can handle stereochemistry and system variation. The consequence is that the model begins with analytical identity of the retention record, not only molecular identity of the analyte.

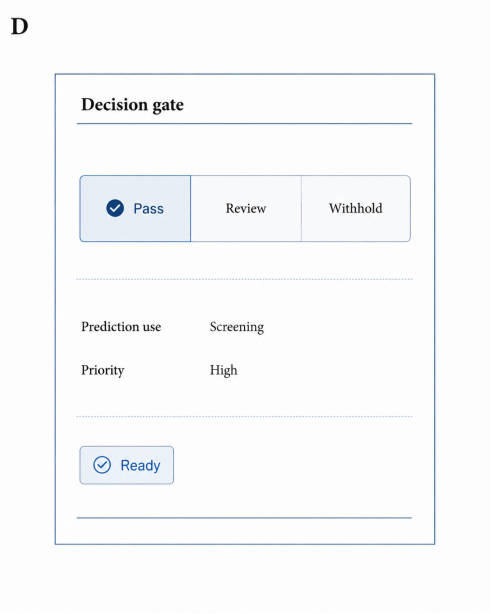
This organization is important because transfer errors often arise before any algorithm is selected. A powerful regression model cannot

Table 3: Validation criteria for reliability gating.

Validation criterion	Interpretive meaning	Role in RG-SC
R-squared	Describes how strongly predicted retention follows observed retention within a defined validation set.	Used as supportive evidence, not as a sole approval criterion.
Mean absolute error	Reports average absolute disagreement between predicted and observed retention values.	Provides an interpretable error summary for analytical users.
Root mean squared prediction error	Emphasizes large prediction errors during external testing.	Identifies local maps that produce unacceptable outliers.
Cross-validation error	Describes internal prediction stability during model selection.	Helps select stable local maps before external evaluation.
Relative mean absolute error	Normalizes absolute error relative to retention magnitude.	Supports comparison across systems with different retention scales.
Relative prediction error	Summarizes proportional prediction disagreement and penalizes large relative deviations.	Penalizes transfer failures when time scales differ across systems.
External predictive ability	Describes performance on compounds withheld from model construction.	Establishes whether the map can support unseen-compound prediction.
Median absolute error	Reports a robust central error under skewed residual distributions.	Protects interpretation when a few compounds cause large deviations.



(a) Applicability check.



(b) Decision gate.

Figure 7: Instrument-facing reliability report.

repair a training set that combines incompatible response types without adequate system labels. RG-SC therefore treats mode assignment, column context, and response type as essential analytical attributes rather than optional metadata. The resulting retention structure supports more cautious transfer decisions: RPLC evidence is allowed to inform RPLC prediction, GC evidence remains retention-index evidence, and mixed records are treated as transfer bridges only when chemical neighbourhood quality supports them.

The design also strengthens compound annotation. In LC–MS procedures, retention estimates can narrow candidate lists produced by accurate mass, isotope pattern, and MS/MS matching, but only if the prediction is traceable to a comparable chromatographic system. In GC–MS procedures, retention-index agreement is most useful when the stationary-phase and temperature-program information are compatible. RG-SC therefore makes chromatographic provenance part of the result: the accepted, advisory, or withheld decision depends on whether the contributing records match the target analytical environment.

3.2. Effect of multi-stream similarity on neighbour selection

The second result is that multi-stream similarity produces a more defensible definition of neighbourhood. Fingerprint similarity alone can overstate relatedness when two compounds share a scaffold but differ in charge state, hydrogen-bonding pattern, or exposed polarity. Distribution-coefficient proximity reduces this risk by incorporating ionization-aware lipophilicity. Retention-factor proximity adds experimental evidence of chromatographic behaviour. Chemical-nature agreement prevents local maps from mixing acids, bases, neutral compounds, zwitterions, and amphiphilic analytes when their retention mechanisms are incompatible.

The added value of second-dominant-interaction compatibility is especially clear for RPLC and HILIC. Hydrophobicity may explain much of RPLC retention, but column selectivity, silanol interactions, dipolarity, hydrogen bonding, and steric shape can determine elution order and peak spacing. The hydrophobic-subtraction literature shows that reversed-phase columns differ in selectivity beyond simple hydrophobic strength [28, 29]. HILIC retention is similarly influenced by polarity, water-layer partitioning, ionic interactions, and hydrogen-bonding capacity. A local map that ignores secondary interactions can appear statistically acceptable while failing in transfer.

Neighbour selection therefore becomes a chemical-instrumental judgement. A query compound is not assigned to the closest structures in an undifferentiated descriptor space; it is assigned to records that are close under the interaction logic of the target separation. This finding directly addresses the paper's research question because transferable retention modelling becomes possible only when local evidence respects both molecular similarity and chromatographic mechanism.

3.3. Reliability-gated prediction and analytical defensibility

The third result is that the reliability gate converts retention prediction into an auditable decision. The gate does not ask only whether a model can produce a numerical value. It asks whether the value is supported by a sufficiently populated local map, compatible chemistry, relevant chromatographic mode, stable validation behaviour, and acceptable error structure. This distinction is important for compound identification, where an incorrect retention-based exclusion can remove the correct candidate from consideration.

The interpretation of validation quantities is also changed. R-squared is treated as a trend descriptor rather than as proof of analytical usefulness. Mean absolute error gives a typical disagreement, but root mean squared prediction error exposes large failures. Relative error is

necessary when retention windows differ across gradients or systems. Median absolute error protects the central interpretation when a few compounds behave unusually. External predictive ability remains the most important test because the intended use is prediction for new compounds.

This gate aligns retention modelling with identification-confidence thinking in analytical chemistry. Accurate mass, isotopic pattern, MS/MS fragments, and retention evidence should not be collapsed into a single unqualified label. Instead, each evidence type should be communicated with its own support and limitations [22, 23]. RG-SC gives retention evidence the same discipline: it can support identification, but only after its local domain has been checked.

3.4. Cross-system transfer

Cross-system transfer is the central stress test for RG-SC. RPLC, HILIC, GC, and mixed systems encode different mechanisms, and a transferable method must keep these systems connected without making them interchangeable. RPLC transfer benefits from records that capture column selectivity and hydrophobic-subtraction behaviour. HILIC transfer requires polarity- and ionization-aware neighbourhoods. GC transfer requires retention-index handling rather than ordinary LC retention-time comparison. Mixed transfer records can connect systems, but only after the gate confirms that chemical neighbourhoods and calibration evidence are suitable.

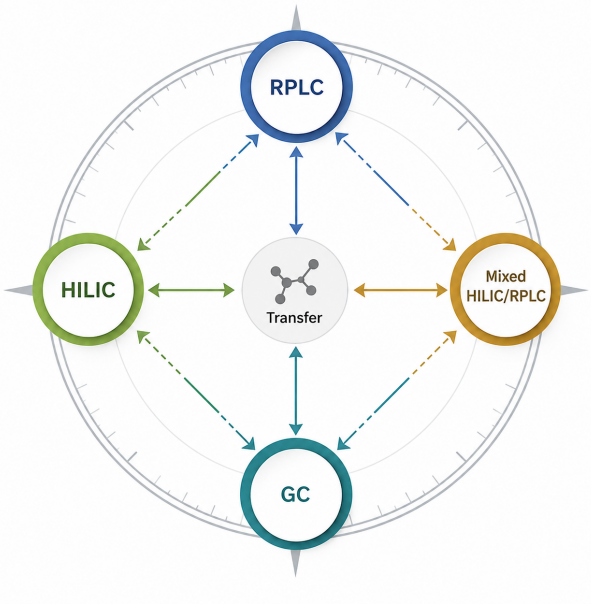


Figure 8: Transfer topology across systems.

The topology in Figure 8 summarizes the transfer result. RPLC, HILIC, and GC are positioned as separate evidence domains linked through a central transfer layer. The arrows do not imply free movement of predictions; they indicate possible transfer routes that must pass through mode-aware calibration and reliability gating. This interpretation prevents a common modelling error: using transfer as a statistical shortcut while ignoring whether the query compound and calibration compounds share the same retention mechanism.

For reversed-phase transfer, METLIN-like, Wilson-type, Tan-type, Hall-type, RIKEN, isocratic, Aicheler-type, and Van Laethem-type records provide several families of local evidence. The method can determine whether a compound lies in a stable RPLC neighbourhood or whether its nearest records come from a column or condition family that should not be transferred. Isocratic evidence is particularly useful when retention-factor behaviour is needed to clarify mechanism under controlled mobile-phase composition.

For HILIC transfer, polar and ionizable compounds require a stricter gate because reversed-phase similarity can be misleading. Retip, Fiehn HILIC, Pathogen Box, and PredRet-type records are useful when they provide neighbours with compatible polarity and chemical nature. When the available neighbours are mostly reversed-phase records, the gate should downgrade or withhold the HILIC prediction even if the structural similarity appears strong. This prevents overconfident transfer into a mode where the dominant interactions have changed.

For GC transfer, retention-index records provide the appropriate response scale. Their value lies in complementing electron-impact spectra and reducing ambiguity in library search, not in providing direct LC retention-time analogues. RG-SC therefore uses GC evidence as a separate retention-index map. This mode-aware handling allows the analytical procedure to operate across LC–MS and GC–MS without confusing their analytical meanings.

3.5. Sparse classes, stereochemistry, and difficult analytical domains

Sparse-domain behaviour is the fourth result. Public retention collections are often dense around common metabolites, pharmaceutical-like compounds, and frequently used chromatographic modes, but they are sparse for rare scaffolds, stereochemical variants, amphiphilic structures, highly ionizable compounds, and forensic toxicology records. Global models can still return predictions for these cases, but such predictions may be unsupported. RG-SC explicitly identifies these domains through neighbourhood density and compatibility checks.

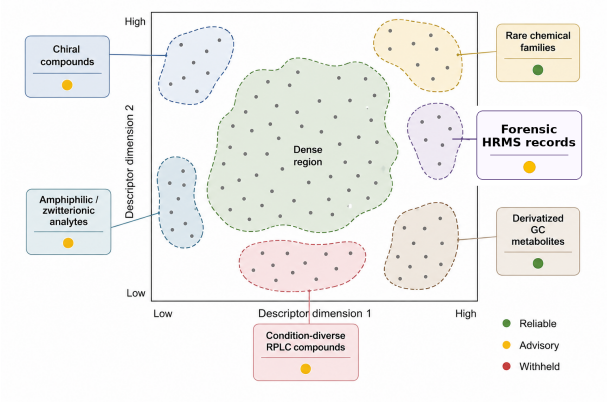


Figure 9: Sparse-domain reliability map.

The sparse-domain map in Figure 9 separates dense, advisory, and withheld regions rather than forcing all query compounds into a single predictive surface. The dense region represents compounds with many compatible neighbours. Advisory regions represent cases where some evidence exists but is incomplete or partly system-dependent. Withheld regions represent cases where chemical class, stereochemistry, or chromatographic condition prevents a defensible prediction. This is not a failure of the model; it is a necessary control for analytical interpretation.

Stereochemical retention illustrates the need for this control. Enantiomers and diastereomers can share most descriptors while showing different retention under chiral or shape-sensitive conditions. A local model that lacks stereochemical evidence should not be used to support chiral retention conclusions. The same logic applies to amphiphilic and highly ionizable analytes, where small changes in pH or ionic strength can shift retention dramatically. In these cases, RG-SC gives analysts a clear message: additional experimental measurement is needed before retention can be used decisively.

Difficult screening domains also benefit from withholding logic. HRMS screening can produce several structurally plausible candidates from accurate mass and fragmentation. Retention can help prioritize

those candidates, but only when comparable records are available. RG-SC therefore protects screening decisions by using retention as a conditional evidence layer rather than a universal filter.

3.6. Retention-assisted candidate filtration

Candidate filtration is the fifth result and the most direct analytical application. A non-targeted LC–MS or GC–MS procedure often produces multiple candidate structures after formula assignment, spectral matching, and database search. Retention evidence can reduce that list by identifying candidates whose chromatographic behaviour is incompatible with the observed window. RG-SC improves this process by applying the reliability gate before retention evidence is used to rank or exclude candidates.

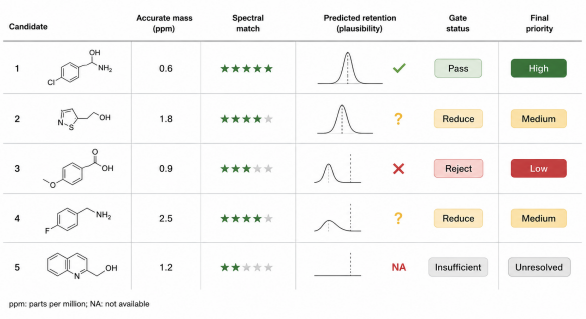


Figure 10: Candidate filtration view.

The filtration view in Figure 10 presents retention as one component of a multi-evidence decision. Accurate mass error, spectral match, predicted retention distribution, gate state, and final priority are kept visible. This layout prevents a retention prediction from being treated as decisive when its gate state is advisory or withheld. It also allows analysts to preserve candidates that have strong spectral evidence but weak retention support, which is important when the retention model lacks local coverage.

This approach is consistent with metabolomics annotation practice, where retention information improves identification confidence but cannot replace authentic standards or high-quality spectral evidence [7, 21, 30]. It also complements retention-order approaches, where relative elution can be more transferable than absolute retention time under certain method changes [31]. RG-SC can incorporate absolute retention time, retention factor, retention order, or retention index depending on the analytical system, provided the reliability gate supports the evidence.

The key operational result is a safer candidate-ranking process. Accepted retention predictions can help prioritize or reject candidates. Advisory predictions can be used as contextual information. Withheld predictions should not influence exclusion decisions. This tiered use of retention evidence reduces false confidence and makes the final annotation more transparent.

3.7. Comparison with conventional QSRR strategies

The final result is a methodological comparison. Conventional global QSRR treats the retention corpus as a single descriptor matrix and usually returns one prediction per query. Fingerprint-only local modelling improves neighbourhood focus but may miss ionization and chromatographic-system mismatch. Deep models can improve representation and scalability, yet they may remain difficult to interpret outside their training domain. RG-SC differs by combining mode separation, multi-stream similarity, validation gating, and the ability to refuse unsupported predictions.

The comparison matrix in Figure 11 shows why RG-SC is not simply another regression model. Its contribution is a retention-decision

Method	Mode separation	Multi-stream similarity	Applicability control	Transfer caution	Candidate filtration readiness	Ability to withhold unsupported prediction
Global QSRR	○	○	◐	◐	◐	○
Fingerprint-only local model	◐	◐	◐	◐	◐	◐
Deep model without gate	◐	◐	○	◐	◐	○
RG-SC (this work)	●	●	●	●	●	●

● Strong ◐ Partial ○ Weak/Absent

Figure 11: Method comparison matrix.

system. Mode separation prevents response-type mixing, multi-stream similarity improves local-neighbour relevance, applicability control limits extrapolation, transfer caution protects cross-system use, and candidate-filtration support connects the model to instrumental practice. The ability to withhold a prediction is especially important because analytical quality control sometimes requires saying that the available evidence is insufficient.

RG-SC can still use many prediction engines inside its local maps. Partial least squares may be appropriate for small interpretable domains. Random forests and gradient boosting may handle nonlinear descriptor relationships. Support-vector regression may be useful in moderate descriptor spaces. Graph neural networks and transformer models may be appropriate when training data are sufficiently large [12–18]. The method is therefore model-agnostic in the prediction step but strict in the evidence and reliability steps.

4. Practical implementation

A practical implementation begins with curation. Molecular structures should be standardized, salts and mixtures should be handled consistently, tautomeric and charge states should be documented, and stereochemistry should be retained when relevant. LC records should preserve chromatographic mode, column chemistry, gradient or isocratic conditions, mobile-phase composition, pH where available, temperature, detection platform, and response type. GC records should preserve retention-index scale, stationary phase, temperature program, derivatization status, and mass-spectral context. Missing metadata should be recorded because it directly affects the gate state.

Descriptor preparation should follow the target mode. RPLC maps should emphasize hydrophobicity, size, ionization, polar surface contribution, and column selectivity. HILIC maps should emphasize polarity, charge behaviour, hydrogen bonding, and zwitterionic character. GC maps should emphasize volatility-related and derivatization-aware descriptors. Chiral maps should include stereochemical descriptors. Descriptor scaling, missing-value treatment, and duplicate-record reconciliation should be documented because they influence neighbour construction.

Local-map construction should proceed in a fixed sequence. The system layer is selected first, then candidate neighbours are screened by fingerprint similarity, distribution-coefficient proximity, retention-factor or retention-index evidence, chemical-nature agreement, and second-dominant-interaction compatibility. A prediction is produced only when the neighbourhood is coherent enough for the selected local model. Reproducibility requires documented machine-learning code, declared software environments, and retained preprocessing records [32]. The output should include the predicted retention value or index, supporting system layer, number and type of neighbours, dominant similarity evidence, validation status, and final gate category.

Reliability communication should be explicit. Accepted predictions can be used for retention-assisted ranking or filtering. Advisory predictions should be communicated as contextual evidence. Withheld predictions should not be used to exclude candidates or claim transfer success. This communication style makes the method suitable for laboratory quality systems because it links model output to the level of action that the evidence can support.

5. Limitations

The main limitation is dependence on metadata quality. Many retention repositories do not provide complete pH, gradient, column, temperature, derivatization, or calibration details. Missing information weakens system compatibility assessment and can move a prediction from accepted to advisory or withheld. Structure standardization is also critical; incorrect tautomer, charge state, salt form, or stereochemistry can place a compound in the wrong neighbourhood.

Sparse domains remain difficult even with a strong reliability gate. RG-SC can prevent overconfident predictions, but it cannot create evidence where no compatible records exist. Rare chemical families, specialized stationary phases, unstable compounds, chiral separations, and highly ionizable analytes may still require experimental retention measurement. GC applications require careful retention-index normalization, while LC applications require detailed method metadata. These limitations define the conditions under which the method should be trusted rather than undermining its purpose.

A further limitation is that the manuscript defines the chemical-retention decision logic rather than giving one pooled numerical comparison across all record families. This choice follows from the research question, which concerns defensible transfer and decision control. Compound-level evaluation should use predefined training, validation, transfer, and external challenge sets so that accepted, advisory, and withheld decisions can be quantified for each chromatographic mode.

6. Conclusion

The research question of this paper was whether heterogeneous chromatographic retention records can be organized into local, mode-aware similarity maps that support transferable QSRR decisions while preventing unsupported predictions from influencing compound identification. The answer is affirmative, provided that transfer is treated as an evidence-gated decision rather than as unrestricted extrapolation. RG-SC answers the question by separating retention evidence into chromatographic-system layers, defining neighbours through six complementary similarity streams, and accepting predictions only after validation behaviour, local density, chemical compatibility, and system relevance are jointly satisfied. The resulting retention corpus is a multi-repository chromatographic evidence system rather than a pooled retention table; the method is a reliability-controlled local mapping procedure rather than a single global regression model; and the result is a decision state that can accept, downgrade, or withhold retention evidence.

The practical conclusion is that retention prediction becomes most useful when its limits are visible. Accepted predictions can support retention-assisted candidate ranking, method transfer, and compound annotation. Advisory predictions can guide interpretation without controlling identification. Withheld predictions protect analysts from false confidence and identify where additional experimental retention measurement is required. This directly serves Chemistry, Analytical, Spectroscopy/Instrumentation/Analytical Sciences because it links molecular structure, chromatographic mechanism, validation, and instrumental decision-making in one auditable analytical procedure. The paper's central contribution is therefore not only a retention-prediction strategy but a reliability discipline for using retention information in LC–MS, GC–MS, metabolomics, environmental analysis, pharmaceutical screening, chiral analysis, and forensic toxicology.

References

[1] Kaliszan, R. (2007). QSRR: Quantitative structure-(chromatographic) retention relationships. *Chemical Reviews*, *107*(7), 3212–3246.

[2] Héberger, K. (2007). Quantitative structure–(chromatographic) retention relationships. *Journal of Chromatography A*, *1158*(1–2), 273–305.

[3] Todeschini, R., & Consonni, V. (2009). *Molecular descriptors for chemoinformatics: Volume I: Alphabetical listing; Volume II: Appendices and references*. John Wiley & Sons.

[4] van den Dool, H., & Kratz, P. D. (1963). A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. *Journal of Chromatography*.

[5] Doley, A., Buragohain, J., & Bharali, A. (2020). Inverse sum indeg status index of graphs and its applications to octane isomers and benzenoid hydrocarbons. *Chemometrics and Intelligent Laboratory Systems*, *203*, 104059.

[6] Domingo-Almenara, X., Guijas, C., Billings, E., Montenegro-Burke, J. R., Uritboonthai, W., Aisporna, A. E., ... & Siuzdak, G. (2019). The METLIN small molecule dataset for machine learning-based retention time prediction. *Nature Communications*, *10*(1), 5811.

[7] Bonini, P., Kind, T., Tsugawa, H., Barupal, D. K., & Fiehn, O. (2020). Retip: Retention time prediction for compound annotation in untargeted metabolomics. *Analytical Chemistry*, *92*(11), 7515–7522.

[8] Stanstrup, J., Neumann, S., & Vrhovsek, U. (2015). PredRet: Prediction of retention time by direct mapping between multiple chromatographic systems. *Analytical Chemistry*, *87*(18), 9421–9428.

[9] Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., & Fiehn, O. (2009). FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry*, *81*(24), 10038–10048.

[10] Cao, M., Fraser, K., Huege, J., Featonby, T., Rasmussen, S., & Jones, C. (2015). Predicting retention time in hydrophilic interaction liquid chromatography mass spectrometry and its use for peak annotation in metabolomics. *Metabolomics*, *11*(3), 696–706.

[11] Rogers, D., & Hahn, M. (2010). Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, *50*(5), 742–754.

[12] Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130.

[13] Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

[14] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232.

[15] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

[16] Yang, Q., Ji, H., Lu, H., & Zhang, Z. (2021). Prediction of liquid chromatographic retention time with graph neural networks to assist in small molecule identification. *Analytical Chemistry*, *93*(4), 2200–2206.

[17] Fedorova, E. S., Matyushin, D. D., Plyushchenko, I. V., Stavrianidi, A. N., & Buryak, A. K. (2022). Deep learning for retention time prediction in reversed-phase liquid chromatography. *Journal of Chromatography A*, *1664*, 462792.

[18] Xue, J., Wang, B., Ji, H., & Li, W. (2024). RT-Transformer: Retention time prediction for metabolite annotation to assist in metabolite identification. *Bioinformatics*, *40*(3), btae084.

[19] Zhang, Y., Liu, F., Li, X. Q., Gao, Y., Li, K. C., & Zhang, Q. H. (2024). Generic and accurate prediction of retention times in liquid chromatography by post-projection calibration. *Communications Chemistry*, *7*(1), 54.

[20] Liu, Y., Yoshizawa, A. C., Ling, Y., & Okuda, S. (2024). Insights into predicting small molecule retention times in liquid chromatography using deep learning. *Journal of Cheminformatics*, *16*(1), 113.

[21] Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., ... & Hankemeier, T. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, *3*(3), 211–221.

[22] Creek, D. J., Dunn, W. B., Fiehn, O., Griffin, J. L., Hall, R. D., Lei, Z., ... & Wolfender, J. L. (2014). Metabolite identification: Are you sure? And how do your peers gauge your confidence? *Metabolomics*, *10*(3), 350–353.

[23] Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying small molecules via high resolution mass spectrometry: Communicating confidence.

[24] Ruttkies, C., Schymanski, E. L., Wolf, S., Hollender, J., & Neumann, S. (2016). MetFrag relaunched: Incorporating strategies beyond in silico fragmentation. *Journal of Cheminformatics*, *8*(1), 3.

[25] Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., ... & Böcker, S. (2019). SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods*, *16*(4), 299–302.

[26] Gramatica, P. (2007). Principles of QSAR model validation: Internal and external. *QSAR & Combinatorial Science*, *26*(5), 694–701.

[27] Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, *29*(6–7), 476–488.

[28] Snyder, L. R., Dolan, J. W., & Carr, P. W. (2004). The hydrophobic-subtraction model of reversed-phase column selectivity. *Journal of Chromatography A*, *1060*(1–2), 77–116.
[29] Wilson, N. S., Nelson, M. D., Dolan, J. W., Snyder, L. R., Wolcott, R. G., & Carr, P. W. (2002). Column selectivity in reversed-phase liquid chromatography: I. A general quantitative relationship. *Journal of Chromatography A*, *961*(2), 171–193.
[30] Garcia, C. A., Gil-de-la-Fuente, A., Barbas, C., & Otero, A. (2022). Probabilistic metabolite annotation using retention time prediction and meta-learned projections. *Journal of Cheminformatics*, *14*(1), 33.
[31] Bach, E., Szedmak, S., Brouard, C., Böcker, S., & Rousu, J. (2018). Liquid-chromatography retention order prediction for metabolite identification. *Bioinformatics*, *34*(17), i875–i883.
[32] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.