Article

GLYBATOMAQ™: Quantum-Geometric Rank Auditing of GLIPR1 Screening by Operator Kernels, DFT/QMC Refinement, and MQWalk Validation

Ioannis Grigoriadis *

Posted Date: 29 May 2026

doi: 10.20944/preprints202605.2096.v1

Keywords: GLIPR1; quantum geometry; positive-semidefinite operator kernels; Bures distance; Fubini-Study metric; quantum Fisher information; DFT/QMC refinement; MQWalk topology validation; rank auditing; virtual screening

doi:10.20944/preprints202605.2096.v1

Article

GLYBATOMAQ™: Quantum-Geometric Rank Auditing of GLIPR1 Screening by Operator Kernels, DFT/QMC Refinement, and MQWalk Validation

Ioannis Grigoriadis [1,2,3,4]

[1] Biogenetoligandorol™ — Department of Computer Drug Discovery Science (Biogenetor™ Stations), Thessaloniki, Greece; jgrigoriadis@biogenea.gr; Tel.:+30 6936592686
[2] GLYBATOMAQ™ Stations, Thessaloniki, Greece
[3] MQWalk™ R&D Units, Thessaloniki, Greece
[4] Biogenea Pharmaceuticals Ltd, Thessaloniki, Greece

Abstract

We present GLYBATOMAQ™, a rank-centric and quantum-geometric framework for GLIPR1-focused in silico screening. The framework treats docking as a fixed-protocol comparative oracle and places the main methodological emphasis on auditable rank movement, positive-semidefinite operator geometry, DFT-derived electronic descriptors, QMC-style uncertainty auditing, and MQWalk topology validation. To make the quantum-geometric contribution explicit for quantum-geometry reporting, we introduce Bures/Fubini-Study-style distance controls, local metric and curvature penalties, quantum Fisher information-inspired sensitivity diagnostics, Berry-type gauge-consistency checks, and a candidate-level assembly certificate. Quantum-geometry reporting elements define the reporting schema, connect each mathematical object to a computational decision, and show how DFT, QMC, MQWalk, curvature, and diagnostic penalties are fused without claiming experimental affinity, efficacy, or biomolecular quantum transport. The output is a reproducible leaderboard and audit bundle for GLIPR1-oriented computational hypotheses: rank shifts are accepted only when supported by electronic descriptors, uncertainty-aware energetic evidence, operator-overlap topology, and chemistry-safe HMC/HSX feasibility constraints.

Keywords: GLIPR1; quantum geometry; positive-semidefinite operator kernels; Bures distance; Fubini-Study metric; quantum Fisher information; DFT/QMC refinement; MQWalk topology validation; rank auditing; virtual screening

---

1. Introduction

1.1. Background and Motivation

Docking-centered virtual screening remains a practical tool for hypothesis generation when applied under fixed, reproducible protocols. Its main utility is comparative prioritization across large libraries rather than literal prediction of binding free energy. This distinction has long been emphasized in the docking literature. AutoDock Vina was introduced as an efficient docking framework for ranking under a defined scoring function [6]. AutoDock4 extended docking with selective receptor flexibility [7]. Glide provided a rapid docking and scoring scheme for pose assessment [8]. At the same time, Kitchen et al. noted that “significant challenges remain” in docking and scoring [9], and Warren et al. presented a critical benchmark-based assessment of docking programs and scoring functions [10]. Later reviews by Pagadala et al. [11] and Guedes et al. [12] likewise stressed that docking scores should be interpreted cautiously. DockThor-VS extends this screening infrastructure, but it does not remove the underlying rank-centric character of the task [13]. Benchmark resources such as PDBbind [14], the docking sets of Huang et al. [15], and DUD-E [16]

doi:10.20944/preprints202605.2096.v1

further reinforce that screening quality is judged primarily through ranking and enrichment behavior. These limitations motivate the introduction of an explicit pattern-recognition layer capable of stabilizing representation, preserving comparability, and making rank behavior auditable. Such a layer is naturally grounded in kernel learning [33], support-vector methods [34], Gaussian-process reasoning [35], and positive-definite matrix analysis [36], with additional numerical support from approximation-theoretic and special-function tools [37-42]. The present study focuses on GLIPR1-oriented in silico screening and is anchored to experimentally determined human sGLIPR1 structures in the Protein Data Bank [1], specifically the structural study of Asojo et al. [2] and the GLIPR1-related entries 3Q2U [3] and 3Q2R [4]. Pocket definition is supported by Fpocket [5]. In this setting, the central problem is not only score generation, but also stable and interpretable ordering of candidates.

1.2. GLYBATOMAQ Concept and the Need for a Chemistry-Safe Consensus

GLYBATOMAQ addresses heterogeneous glioma-associated ligand collections through a hierarchical Hyper Mega Core (HMC) concept that is intentionally chemistry-safe. Rather than enforcing atom-level fusion across non-comparable chemotypes, the framework defines consensus at the level of family-aware pharmacophore grammar and modular scaffold logic. This preserves medicinal-chemistry realism while enabling descriptor-space comparison and rank-based evaluation across diverse ligand families. This design is consistent with both classical and modern representational chemistry. Practical curation can be carried out with RDKit [18], Morgan-style structure encoding [19], extended-connectivity fingerprints [20], and SMILES-based normalization [21]. SELFIES provides a robust molecular string representation [22]. More structured generative and graph-based formulations, including grammar VAE [23], junction-tree VAE [24], graph-policy learning [25], recurrent molecular design [28], latent continuous molecular design [29], neural message passing [30], graph convolution [31], and attention-based architectures [32], all support the same broader principle: chemically meaningful comparison depends on structured validity and interpretable abstraction rather than indiscriminate atom-level merging. Descriptor-space proximity, therefore, is not taken to imply synthetic mergeability. In GLYBATOMAQ, consensus representations are used for comparison, clustering, and ranking, not as synthetic templates.

1.3. GLYBATOMAQ™ as a Quantum-Enriched Ranking Extension

GLYBATOMAQ™ extends the base GLYBATOMAQ framework through a late-stage, quantum-enriched refinement layer positioned after docking-oriented pattern recognition and before final prioritization. The extension is deliberately conservative. Docking remains the fixed-protocol oracle [6-8,13], while pattern recognition remains the principal ranking machinery [33-36]. The added DFT, QMC, and MQWalk components are used for re-ranking, uncertainty quantification, and topology validation rather than for efficacy or clinical inference. The term "quantum-enriched" is methodological, not physical. It denotes a refinement layer that supplements an existing rank frontier with deterministic electronic-structure cues, stochastic uncertainty-aware auditing, and graph-propagation-based topology checks. The DFT component is grounded in the Hohenberg–Kohn [44] and Kohn–Sham [45] framework, with standard interpretive support from Parr and Yang [46] and Martin et al. [47]. The QMC component follows the uncertainty-aware logic of Foulkes et al. [48] and Needs et al. [49]. MQWalk uses notation consistent with quantum-information formalisms [50], but its interpretation remains strictly algorithmic in view of broader cautions concerning extrapolation of quantum language to biology [51,52]. The workflow is therefore best understood as a rank augmentation scheme: Dock/Rank → Shortlist → DFT/QMC/MQWalk → Fuse → Re-rank + Audit. This interpretation is also relevant to Pattern Recognition. Recent work has emphasized ranking-based learning [56], local-kernel graph learning [57], informative graph selection [58], dynamic graph representation [59], adaptive graph propagation [60], and molecular graph contrastive learning [61]. GLYBATOMAQ™ adopts the same broad principle in a molecular screening setting: reliable decisions on structured data require stable representation, explicit propagation logic, and interpretable uncertainty control.

doi:10.20944/preprints202605.2096.v1

1.4. Core Technical Contributions

GLYBATOMAQ™ introduces three complementary refinement components. First, the DFT branch computes compact electronic summaries for shortlisted candidates under a fixed protocol, including energy proxies, frontier-gap terms, charge or dipole statistics, and self-consistent-field diagnostics, with reliability-aware scoring and explicit failure logging [44-47]. Its role is not to replace docking globally, but to resolve crowded near-tie regions in a controlled way. Second, the QMC branch is applied only to a smaller subset of candidates and provides energy-like estimates together with explicit uncertainty summaries, including variance, confidence-interval width, and effective sample size [48,49]. In this sense, QMC functions as a selective audit mechanism rather than a screening-scale rescoring engine. Third, Avogadro-conditioned MQWalk topology validation reuses the normalized operator-overlap graph to test whether DFT- and QMC-driven promotions remain compatible with the topology implied by the representation layer. Its operator-kernel basis is grounded in kernel methods [33], positive-definite matrix structure [36], and bounded analytic conditioning [37-43], whereas the walk notation is borrowed from formal propagation language [50]. Here again, the interpretation is strictly algorithmic: MQWalk is topology-sensitive, not a claim of biomolecular quantum transport [51,52]. 1.5. Scope boundary and interpretation All outputs generated by GLYBATOMAQ™ should be interpreted as computational hypotheses for prioritization and analysis. The framework is intended to improve rank stability, auditability of candidate movement, uncertainty-aware re-ranking, and visibility of failure modes while remaining fully compatible with docking-centered virtual screening. It does not establish binding affinity, mechanism of action, pharmacology, safety, or clinical efficacy [9-12,14-16]. Accordingly, GLYBATOMAQ™ is best positioned alongside, rather than against, contemporary machine-learning-guided molecular-modeling workflows. The broader ecosystem includes structured generative chemistry [23-25,28,29], message-passing and graph-convolutional learning [30,31], attention-based architectures [32], AlphaFold [53], AlphaFold 3 [54], and DiffDock [55]. These advances expand the front end of molecular modeling, but they do not eliminate the downstream need for careful rank interpretation, reproducible protocol boundaries, and post hoc auditing of candidate movement. GLYBATOMAQ™ therefore occupies a narrower role: it is a reproducible post-docking ranking-and-analysis framework, not a substitute for experiment and not a basis for direct therapeutic inference.

2. Methods and Materials

2.1. GLYBATOMAQ™ Quantum-Enriched Re-Ranking Layer: Mathematically Audited DFT/QMC Refinement with MQWalk Topology Validation

2.1.1. Scope, Rationale, and Methodological Placement Within a Rank-Centric GLYBATOMAQ Workflow

GLYBATOMAQ™ is formulated as a late-stage, rank-centric mathematical-chemistry re-ranking layer rather than as a de novo docking, binding-free-energy, or generative-design engine. Its methodological purpose is deliberately narrow: to improve the interpretability and reliability of a pre-existing molecular ranking frontier by combining deterministic electronic descriptors, stochastic energetic uncertainty, and operator-kernel topology diagnostics. This placement is important because the proposed contribution is not the invention of a new docking score. Instead, GLYBATOMAQ™ addresses a recurring mathematical-chemistry problem: how to decide whether a ligand promoted by a screening pipeline is supported simultaneously by local electronic structure, uncertainty-aware energetic evidence, and global similarity topology.

Docking is therefore treated as a fixed-protocol comparative signal, not as an experimental surrogate. AutoDock Vina [6], AutoDock4 [7], Glide [8], and DockThor-VS [13] are used only to define an upstream ranking frontier. This interpretation follows established cautions in the docking literature. Kitchen et al. noted that “significant challenges remain” in docking and scoring [9], Warren et al. critically assessed scoring-function behavior [10], and subsequent reviews likewise emphasized docking as a comparative prioritization tool rather than a literal affinity estimator [11,12]. Benchmark context is further provided by PDBbind [14], the docking test sets of Huang et al. [15], and DUD-E [16]. Thus, the first methodological safeguard is conceptual: docking creates a hypothesis-generating frontier, whereas GLYBATOMAQ™ audits rank movement within that frontier.

The structural context is introduced through the Protein Data Bank [1], the GLIPR1 structural study of Asojo et al. [2], and the GLIPR1-related structures 3Q2U [3] and 3Q2R [4], with pocket localization supported by Fpocket [5]. These structures define the protein-side objects on which the downstream rank audit operates. Importantly, the protein pockets are not treated as interchangeable docking boxes. They are treated as distinct local chemical environments, each inducing its own ligand ordering, operator representation, and refinement uncertainty. This pocket-conditioned design is essential because mathematical-chemistry rank stability is rarely global: a ligand may be plausible in one structural microenvironment and weakly supported in another.

The re-ranking layer is invoked only for a pre-ranked shortlist and resolves near ties through three coupled branches: DFT for deterministic electronic descriptors [44-47], QMC for stochastic energetic auditing with explicit uncertainty [48,49], and MQWalk for topology validation against the upstream positive-semidefinite operator-kernel similarity space [33,36,50]. The novelty of the layer lies in this constrained coupling. DFT contributes local electronic information; QMC contributes uncertainty-aware energetic stress testing; MQWalk contributes a graph-theoretic consistency check on whether a rank promotion is compatible with the ligand–pocket similarity geometry that originally placed the candidate in the shortlist. The method therefore converts a potentially opaque score change into a decomposable rank-audit event.

The term “quantum walk” is used strictly as an algorithmic graph-propagation analogy and not as a claim that biomolecular recognition is mediated by coherent quantum transport. This distinction is explicit to avoid overinterpretation and is consistent with the cautions of Arndt et al. [51] and Nunn et al. [52]. In GLYBATOMAQ™, MQWalk is a controlled mathematical diagnostic on a normalized operator graph. It is used to ask whether rank promotions remain stable under unitary-style propagation on the similarity network, not to claim a physical quantum walk inside GLIPR1.

Operationally, the workflow is:

The supporting visual evidence is cited rather than reproduced in the main manuscript: SI Appendix I Figure 1 supplies the MMP3 inhibitor rank/MD benchmark context, SI Appendix I Figure 2 supplies the vortioxetine glioblastoma network-pharmacology and docking numerics, SI Appendix I Figs. 3-6 document the HSX dashboard, persistence heatmaps, DFT-branch diagnostics, MQWalk topology validation, and residue-energy fingerprints, and SI Appendix I Figs. 7-9 provide the AutoDock Vina and DockThor leaderboard audit trails. These supplementary figures are used as quantitative context for the main Figures 10-12 and Tables 12-15 without transferring the supplementary panels into the main article.

Dock/Rank → Shortlist → DFT/QMC/MQWalk → Fuse → Re-rank + Audit

as illustrated in SI Appendix I Figure 1A and SI Appendix I Figure 2. The returned object is not a single scalar score but a rank-audit bundle containing $S_{\mathrm{DFT}}(P,L)$, $S_{\mathrm{QMC}}(P,L)$, $S_{\mathrm{MQW}}(P,L)$, reliability terms, diagnostic penalties, and the post-refinement rank shift $\Delta R(P,L)$. This distinction is central to the research-level mathematical-chemistry value of the method: the output is a structured explanation of ranking evidence rather than a black-box replacement score. The formulation is naturally aligned with Pattern Recognition because the primary object is a structured ranking problem, consistent with ranking-based transformation [56], local-kernel graph learning [57], informative graph selection [58], dynamic graph representation [59], adaptive graph propagation [60], and molecular graph contrastive learning [61].

2.1.2. Inputs, Data Model, and Candidate-Selection Boundary

The GLYBATOMAQ™ refinement layer consumes only protocol-logged upstream outputs. These include pocket definitions P anchored to 3Q2U [3] and 3Q2R [4], optionally cross-checked by Fpocket [5]; a ligand set L curated under chemistry-safe provenance; docking outputs $D(P,L)$

generated under fixed settings; and the base GLYBATOMAQ score $S_\theta(P,L)$, which defines the pre-refinement frontier R_0 while remaining distinct from docking. This separation prevents circular interpretation: docking contributes poses and an initial comparative frontier, whereas S_θ defines the model-derived order to be audited.

The chemistry-safe representation layer is compatible with MMFF94 geometry handling [17], RDKit [18], Morgan-style encodings [19], ECFP fingerprints [20], SMILES [21], and SELFIES [22]. It is also compatible with structured molecular learning approaches including grammar VAE [23], junction-tree VAE [24], graph-policy learning [25], recurrent molecular design [28], latent continuous design [29], neural message passing [30], graph convolution [31], and attention-based architectures [32]. These representations are not introduced as independent claims of predictive superiority. Rather, they define admissible molecular encodings from which stable candidate objects and operator summaries can be constructed.

At the operator level, pocket and ligand topology inputs are summarized by positive-semidefinite matrices ρ_P and ρ_L. A typical construction is

$$ϱX = nX^{-1}\, X\, X^{T},\ ϱX \succeq 0,\ X \in P,L \tag{1}$$

where Z is a logged feature or embedding matrix after preprocessing, n is the number of encoded objects or local descriptors, and ρ is therefore a Gram-type positive-semidefinite operator. The ligand–pocket similarity is then expressed through the normalized overlap kernel

$$\bar{K}(P,L) = \mathrm{Tr}(ϱP\ ϱL) / \sqrt{}(\mathrm{Tr}(ϱP^2)\ \mathrm{Tr}(ϱL^2) + \varepsilon) \tag{2}$$

This quantity has two roles. First, it provides a scale-controlled similarity measure between protein-pocket and ligand operator summaries. Second, it supplies the graph weights used by MQWalk validation. Because $\bar{K}(P,L)$ is derived from positive-semidefinite objects, the resulting geometry is compatible with kernel learning [33], Gaussian-process similarity concepts [35], and positive-definite matrix theory [36]. This operator-overlap construction is one of the main mathematical-chemistry innovation clues in the workflow: it replaces an isolated ligand score with a trace-normalized comparison between pocket and ligand chemical information operators.

The role of this stabilizing layer between docking and refinement is shown in SI Appendix I Figure 2 and recurs in SI Appendix I Figure 6. In practical terms, a ligand can be promoted after DFT/QMC refinement only if the promotion is not discordant with the underlying operator-kernel topology. This prevents the refinement layer from acting as an unconstrained score amplifier.

To control computational cost and to preserve the intended statistical meaning of the procedure, refinement is applied only to shortlist candidates defined by

$$L_{top}(P) = L \in LP : \mathrm{rank}(S\theta(P,L)) \le K_{short} \tag{3}$$

When required, a consensus shortlist is pooled across pockets. Each evaluated pair $\left(P^{(m)},L\right)$ is stored as a candidate object

$$O(P,L) = P,\ L,\ \chi L,\ D(P,L),\ S\theta(P,L),\ ϱP,\ ϱL,\ \bar{K}(P,L),\ \Pi_{prov} \tag{4}$$

where $\mathrm{Conf}(L) = \{\chi_c\}$ denotes the conformer ensemble and Π_{prov} records provenance, descriptor settings, conformer source, docking protocol, and refinement status. The purpose of this object is replayability. Any rank movement must be traceable to a defined input, descriptor, uncertainty estimate, or topology diagnostic. The shortlist-and-branch logic is illustrated in SI Appendix I Figure 1A, while the chemistry-safe HSX anchor view is visible in SI Appendix I Figure 2 and SI Appendix I Figure 3.

2.1.3. Shared Conditioning, Robust Scaling, and Avogadro-Anchored Numerical Regularization

Before branch scoring, all candidate features are subjected to shared conditioning. For each feature channel f_k, a robust standardized value is computed as

$$z_j = (x_j - \mathrm{median}(x_j)) / (\mathrm{MAD}(x_j) + \varepsilon) \tag{5}$$

This choice is used because shortlist distributions are typically small, heavy-tailed, and sensitive to outliers. A mean–variance normalization would allow a single unstable electronic or docking-derived channel to dominate the branch fusion. Robust normalization therefore serves a mathematical role: it converts heterogeneous chemical descriptors into bounded, comparable evidence components.

The standardized channel is then scaled by an Avogadro-anchored numerical factor

$$aN = \exp(-\alpha \log NA) = NA^{-}\alpha \tag{6}$$

and bounded as

$$\hat{x_j} = \tanh(aN\ z_j) \tag{7}$$

where N_A is the CODATA Avogadro constant used only as a deterministic scale anchor [43]. This step is numerical rather than physical. No molecular-counting interpretation is assigned to N_A. The reason for using this anchor is to provide a reproducible, dimension-aware damping convention in a workflow that combines graph, electronic, and stochastic energetic quantities. The damping prevents branch-level descriptors with different native scales from dominating the fused rank score.

This conditioning step is consistent with the manuscript's kernel and approximation-theoretic basis [33,36-42]. Its placement before branch fusion is depicted in SI Appendix I Figure 2 and SI Appendix I Figure 5B. Methodologically, the step should be read as a reproducibility safeguard: all branch scores are built from traceable, robustly scaled, and bounded quantities before they are allowed to influence rank movement.

2.1.4. DFT Branch: Deterministic Electronic-Structure Descriptors for Near-Tie Separation

The DFT branch follows the Hohenberg–Kohn [44] and Kohn–Sham [45] formalism, interpreted in the standard DFT treatments of Parr and Yang [46] and Martin et al. [47]. Its function in the workflow is not to produce an absolute binding free energy. Instead, DFT supplies deterministic electronic-structure descriptors for shortlist candidates whose upstream rankings are too close to be resolved confidently by docking and operator-kernel information alone.

For a logged geometry $\chi^{\star}(P, L)$, the formal descriptor basis is the Kohn–Sham relation

$$[-1/2\ \nabla^2 + v_{eff}[\varrho](r)]\ \psi_i(r) = \varepsilon_i\ \psi_i(r) \tag{8}$$

together with the variational condition

$$\delta EKS[\varrho]\ /\ \delta\varrho(r) = \mu \tag{9}$$

From this calculation, the branch extracts a descriptor vector

$$\Phi DFT = (EKS,\ \Delta\varepsilon HL,\ \mu_e,\ \eta_e,\ \omega_e,\ q_{res},\ \mu^{d}_{ip},\ cSCF) \tag{10}$$

where the specific components are selected from logged electronic descriptors, frontier-orbital summaries, charge descriptors, local contact descriptors, and density-derived quantities. The exact descriptor list is fixed before evaluation to avoid post hoc selection.

The DFT branch score is written as

$$SDFT(P,L) = \beta DFT^{T}\ \Phi DFT(P,L) \tag{11}$$

with a reliability gate

$$rDFT = ISCF\ I_{geom}\ I^{d}_{omain}\ I_{complete} \tag{12}$$

The reliability term records convergence quality, geometry stability, descriptor completeness, and whether the computed electronic descriptors fall within the calibration envelope of the shortlist. Thus, DFT contributes deterministic near-tie separation only when the calculation is internally stable. If convergence is poor or the electronic descriptors are outside the expected numerical domain, the DFT contribution is down-weighted rather than silently retained.

This branch is visualized most directly in the descriptor histograms and scatter diagnostics of SI Appendix I Figure 5A. Its mathematical-chemistry contribution is the conversion of local electronic structure into a controlled, reliability-gated rank perturbation. This is distinct from using DFT as an unqualified scoring replacement.

doi:10.20944/preprints202605.2096.v1

2.1.5. QMC Branch: Uncertainty-Explicit Energetic Auditing

The QMC branch follows the uncertainty-aware logic of Foulkes et al. [48] and Needs et al. [49]. Its purpose is selective stochastic auditing rather than screening-scale rescoring. The branch is applied to candidates for which the rank frontier is sensitive to electronic-energy uncertainty, DFT-derived separation is small, or topology diagnostics identify a candidate whose promotion requires additional energetic scrutiny.

For each audited pair (P, L), the QMC branch produces an estimated energy

$$E\_QMC = M^{-1} \Sigma_{m=1}^{M} EL(R_m) \tag{13}$$

an uncertainty estimate

$$\sigma\_QMC^2 = (M-1)^{-1} \Sigma_m [EL(R_m) - E\_QMC]^2 \tag{14}$$

and a confidence-width summary

$$CI\_QMC = z_1 - \alpha/2\ \sigma\_QMC / \sqrt{}M_{e}\not{f}\not{f} \tag{15}$$

The conservative QMC score is defined as

$$SQMC = -E\_QMC - \lambda\sigma\ \sigma\_QMC - \lambda CI\ CI\_QMC \tag{16}$$

This expression deliberately penalizes uncertain energetic estimates. A candidate is not rewarded merely for having a favorable stochastic estimate if that estimate is accompanied by a large uncertainty interval. The QMC branch therefore introduces a research-level uncertainty principle into the ranking procedure: rank promotion must survive both the expected energetic value and the uncertainty attached to that value.

The QMC contribution is gated by

$$rQMC = I_{sample}\ I_{var}\ ICI\ I_{conv} \tag{17}$$

which records sampling adequacy, variance behavior, confidence-width stability, and convergence diagnostics. Candidates with inflated uncertainty are flagged rather than simply demoted, because uncertainty inflation is itself chemically informative. It may indicate conformational ambiguity, sensitivity to local geometry, or instability in the electronic-energy estimate. The QMC branch is shown as an uncertainty-explicit component in SI Appendix I Figure 2, summarized in SI Appendix I Figure 3, and fused in SI Appendix I Figure 5B.

2.1.6. MQWalk Branch: Operator-Kernel Topology Validation of Rank Promotions

MQWalk is the topology-validation branch of GLYBATOMAQ™. Its purpose is to determine whether DFT/QMC-supported rank movement remains consistent with the ligand–pocket similarity geometry induced by the normalized operator-overlap kernel. Unlike DFT and QMC, MQWalk does not add a physical electronic-energy model. It is a mathematical diagnostic on the shortlist graph.

Starting from the normalized overlap kernel $\bar{K}(P, L)$, a bipartite pocket–ligand adjacency matrix is defined as

$$APL = \bar{K}(P,L) \tag{18}$$

The degree-normalized adjacency is

$$AN = D^{-1}/2\ A\ D^{-1}/2 \tag{19}$$

where D is the diagonal degree matrix. To prevent disconnected or weakly connected shortlist components from producing unstable propagation behavior, teleport mixing is introduced:

$$A\gamma = (1-\gamma)\ AN + \gamma\ 11^T/|V| \tag{20}$$

The Avogadro-conditioned graph Laplacian is then

$$\tilde{L} = NA^{-}\alpha(I - A\gamma) \tag{21}$$

Propagation is written as

$$U(t) = \exp(-i\ t\ \tilde{L}) \tag{22}$$

and the node mass at node j is

$$p_j(t) = |\langle j|U(t)|\psi_0\rangle|^2 \tag{23}$$

These quantities are interpreted only as topology-consistency diagnostics. The central question is whether a promoted ligand remains close, under graph propagation, to the pocket and ligand neighborhoods that justified its initial ranking. If a candidate receives a strong DFT or QMC score but is topologically isolated from chemically similar pocket–ligand neighborhoods, the promotion is marked as discordant.

The MQWalk score is defined as

$$\mathrm{SMQW} = \lambda \mathrm{w}_{\mathrm{alk}} \langle \mathrm{p(t)} \rangle_{\mathrm{path}} - \lambda^{\mathrm{d}}{}_{\mathrm{isc}}\, \Delta_{\mathrm{topo}} \tag{24}$$

Here, MQWscore (P, L) summarizes propagation-supported proximity, while $\Delta_{\mathrm{topo}}(P, L)$ penalizes discordance between branch-level promotion and graph-level neighborhood consistency. In practice, MQWalk functions as a graph-consistency filter rather than as an additional physical model. It is specifically designed to prevent unsupported rank jumps caused by a single favorable branch score.

This branch is represented in SI Appendix I Figure 1B–C, SI Appendix I Figure 3, and SI Appendix I Figure 6. Its main mathematical-chemistry innovation is that ligand prioritization is constrained not only by scalar descriptors but also by the topology of a positive-semidefinite ligand–pocket operator space.

2.1.7. Rank Movement, Diagnostic Penalties, and Reliability-Aware Fusion

Rank movement is audited explicitly. For each pocket–ligand pair, the pre-refinement rank is $R_0(P, L)$, the post-refinement rank is $R_+(P, L)$, and the rank shift is

$$\Delta \mathrm{R(P,L)} = \mathrm{R_0(P,L)} - \mathrm{R^+(P,L)} \tag{25}$$

A positive ΔR indicates promotion, while a negative value indicates demotion. Rank stability is summarized by

$$\mathrm{Overlap@k} = |\mathrm{Top_k(R_0)} \cap \mathrm{Top_k(R^+)}| \,/\, \mathrm{k} \tag{26}$$

This statistic is included because a useful refinement layer should not arbitrarily destroy the upstream frontier. A low Overlap@k value is not automatically wrong, but it requires explanation by DFT evidence, QMC evidence, or topology-supported correction. Thus, rank movement is treated as an object requiring justification.

Diagnostic penalties are collected in

$$\Pi\_\mathrm{diag} = \lambda \mathrm{D}({}_1 - {}_\mathrm{r}\mathrm{DFT}) + \lambda \mathrm{Q}({}_1 - {}_\mathrm{r}\mathrm{QMC}) + \lambda \mathrm{T}\, \Delta_{\mathrm{topo}} + \lambda \mathrm{M}\, \mathrm{I}_{\mathrm{missin9}} \tag{27}$$

which includes flags for low DFT reliability, inflated QMC uncertainty, convergence failure, descriptor-domain violation, topology discordance, or missing provenance. These flags are not cosmetic. They determine whether a candidate's final rank is considered supported, weakly supported, or audit-failed.

The final fused score is

$$\mathrm{S^+(P,L)} = \mathrm{w_0\, S\theta + wD\, rDFT\, SDFT + wQ\, rQMC\, SQMC + wM\, SMQW} - \Pi\_\mathrm{diag} \tag{28}$$

All weights are predeclared or logged for replayability. The fusion rule is therefore not an opaque ensemble. It is a reliability-aware evidence-integration rule in which each branch has a distinct mathematical role:

1. S_θ preserves the upstream ranking frontier.
2. S_{DFT} introduces deterministic electronic-structure separation.
3. S_{QMC} introduces uncertainty-penalized energetic auditing.
4. S_{MQW} constrains rank movement by operator-kernel topology.
5. Π_{diag} prevents unreliable branch evidence from being overinterpreted.

The full fused logic is summarized in SI Appendix I Figure 5B, while SI Appendix I Figure 6 shows how hotspot, topology, and residue-level evidence can be interpreted as a unified audit trail. This audit-first design is the principal methodological answer to the possibility that the workflow may appear uncontrolledly complicated. Each component is retained only if it performs a nonredundant role in the rank-audit chain.

2.1.8. Pocket-Specific and Consensus Leaderboards

Because pocket-dependent reversals are expected, the method reports both pocket-specific leaderboards and a consensus leaderboard. For each pocket P, the pocket-specific leaderboard is obtained by sorting $S_{\mathrm{final}}(P,L)$. This preserves the possibility that different GLIPR1 pocket environments support different ligand hypotheses.

The consensus score is defined as

$$S_{\mathrm{cons}}(L) = \mathrm{Agg}P \in \mathcal{P}\ S^{+}(P,L) \qquad (29)$$

where the aggregation rule is predeclared and logged. The aggregation may be a mean, trimmed mean, rank-product, Borda-type statistic, or another fixed rank-aggregation rule, but it must be specified before evaluating the final leaderboard. The consensus result is therefore not a post hoc selection of the most favorable pocket result. It is a reproducible summary of pocket-conditioned evidence.

This reporting strategy is consistent with benchmark-aware docking practice discussed by Kitchen et al. [9], Warren et al. [10], Pagadala et al. [11], Guedes et al. [12], Huang et al. [15], and Mysinger et al. [16], and with recent Pattern Recognition work emphasizing ranking, graph consistency, and structured aggregation rather than isolated scalar decisions [56-61]. In practical terms, the pocket-wise and consensus views implement the manuscript's broader audit-first philosophy: docking defines the initial frontier, refinement resolves near ties, topology checks constrain unsupported jumps, uncertainty terms penalize unstable energetic evidence, and the final output is a replayable leaderboard rather than an opaque score list. This interpretation is reinforced across SI Appendix I Figure 1, SI Appendix I Figure 2, SI Appendix I Figure 5, SI Appendix I Figure 6, and the docking-led leaderboard summaries in SI Appendix I Figure 7–9.

2.1.9. Mathematical-Chemistry Innovation, Integrated DFT/QMC Refinement, MQWalk Topology Validation, and Methodological Limits

The methodological innovation of GLYBATOMAQ™ is not merely that it combines docking, DFT, QMC, graph kernels, and quantum-walk-style propagation in one workflow. Such a combination would be insufficient as a mathematical-chemistry contribution if the components were only assembled sequentially or used as non-functional computational layers. The specific contribution is instead the construction of a **rank-audit formalism** in which docking, operator-kernel geometry, electronic refinement, stochastic uncertainty, and topology validation play distinct, nonredundant roles. The workflow is therefore best summarized as a constrained evidence chain,

$$\text{Dock/Rank} \rightarrow \text{PSD operator geometry} \rightarrow \text{DFT/QMC audit} \rightarrow \text{MQWalk validation} \rightarrow S^{+} \rightarrow \Delta R \qquad (30)$$

This chain is the central mathematical-chemistry object of GLYBATOMAQ™. Docking defines an upstream frontier; positive-semidefinite operator kernels impose a reproducible geometry on ligand–pocket comparisons; DFT and QMC supply electronic-structure and uncertainty-aware refinement; MQWalk checks whether rank promotions remain compatible with the learned topology; and the final output is a replayable leaderboard with diagnostic penalties rather than an opaque scalar claim. This logic is illustrated globally in Figure 1A, Figure 2, Figure 5B, and Figure 6, and it is further instantiated in the docking-led leaderboard views of Figure 7, Figure 8, and Figure 9.

The methodological boundary is deliberately conservative. GLYBATOMAQ™ does not claim that docking scores are binding affinities, that DFT or QMC rescoring replaces experimental binding measurements, or that MQWalk describes physical quantum transport in biomolecular recognition. Instead, the method addresses a narrower and more defensible mathematical-chemistry problem: **when a ligand is promoted or demoted after refinement, can that rank movement be decomposed**

into electronic, energetic, uncertainty, and topology-consistency evidence? The value of the method lies in answering that question explicitly.

2.1.9.1. Ranking, not Absolute Affinity, as the Primary Mathematical Object

The first innovation is the treatment of ligand prioritization as a ranking problem rather than as an isolated score-regression problem. In the GLYBATOMAQ™ setting, the initial docking output is written as

$$D(P,L,c) = \text{score}^{d}{}_{\text{ock}}(P,L,c) \tag{31}$$

where P denotes a pocket, L denotes a ligand, and χ denotes a ligand conformer or docked pose. This score is used only as a fixed-protocol oracle under a defined docking setup. It is not interpreted as a direct experimental affinity. This interpretation is consistent with the known limitations of docking and scoring functions [6-16].

The upstream GLYBATOMAQ score $S_\theta(P,L)$ defines a pre-refinement rank,

$$\Delta R(P,L) = R_0(P,L)-R^{+}(P,L) \tag{32}$$

After DFT/QMC/MQWalk refinement and reliability-aware fusion, the final score $S_{\text{final}}(P,L)$ defines a post-refinement rank,

$$R^{+}(P,L) = \text{rank}L\in L_{\text{top}}(P)[-S^{+}(P,L)] \tag{33}$$

The central reported quantity is the rank shift,

$$\Delta R(P,L) = R_0(P,L) - R^{+}(P,L) \tag{34}$$

A positive value indicates promotion after refinement, whereas a negative value indicates demotion. The method therefore does not ask only whether a ligand has a numerically favorable score. It asks whether that ligand's **rank movement** is supported by chemically interpretable and mathematically stable evidence.

This distinction is especially important in near-tie regions. If two ligands differ by a very small docking-score interval, a naive workflow may overinterpret that difference. GLYBATOMAQ™ instead treats such cases as requiring additional branch evidence. For a best-pocket docking summary,

$$D_{\beta\text{est}}(L) = \text{min}P,c\ D(P,L,c) \tag{35}$$

a distance-to-best statistic may be written as

$$\delta_{\beta\text{est}}(L) = D_{\beta\text{est}}(L) - \text{min}L'D_{\beta\text{est}}(L') \tag{36}$$

This statistic is used in Figure 7 and Figure 8 to distinguish strong separation from near-tie behavior. In the extracted AutoDock Vina leaderboard, GLYBATOMAQ-HSX-CONS2 and GLYBATOMAQ-HSX-cons4 differ by only approximately 0.1 kcal/mol in their best-pocket values, whereas larger gaps separate them from several reference agents. In the GLYBATOMAQ™ interpretation, the small separation between the leading HSX candidates is not treated as a final biochemical conclusion. Instead, it identifies a region where DFT/QMC refinement and MQWalk topology validation are most useful.

Rank stability is further summarized by overlap statistics,

$$\Omega_k = |\text{Top}_k(R_0) \cap \text{Top}_k(R^{+})|/k \tag{37}$$

A high value indicates that refinement preserves the upstream frontier, whereas a low value indicates major reordering. Major reordering is not automatically undesirable, but it must be explained by branch-level evidence. This is why the GLYBATOMAQ™ output is not simply S_{final}. It is a rank-audit object containing S_{DFT}, S_{QMC}, S_{MQW}, reliability gates, diagnostic penalties, ΔR, and Overlap@k . This rank-centric interpretation is consistent with recent Pattern Recognition work emphasizing ranking transformations, graph learning, adaptive propagation, and molecular graph representation [56-61].

2.1.9.2. Positive-Semidefinite Operator Geometry as the Stabilizing Mathematical Layer

The second innovation is the embedding of pocket and ligand information into positive-semidefinite operator summaries. Instead of comparing ligands and pockets only through scalar docking scores, GLYBATOMAQ™ constructs operator-valued summaries,

$$\varrho P,\ \varrho L \in S^{+d},\ \mathrm{Tr}(\varrho P) = \mathrm{Tr}(\varrho L) = 1 \tag{38}$$

A typical construction is

$$\varrho X = ZX^{T}\, ZX \,/\, \mathrm{Tr}(ZX^{T}\, ZX) \tag{39}$$

where Z is a logged feature, descriptor, or embedding matrix derived from chemistry-safe molecular encodings, pocket descriptors, interaction fingerprints, or learned representations. This construction is compatible with RDKit, MMFF94, Morgan/ECFP-style fingerprints, SMILES, SELFIES, graph neural encodings, attention-based encodings, and molecular graph-learning representations [17-32].

The ligand–pocket operator overlap is written as

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{40}$$

and its normalized form is

$$\bar{K}(P,L) = K(P,L)/(\lVert \varrho P \rVert F\ \lVert \varrho L \rVert F + \varepsilon) \tag{41}$$

This normalized trace-overlap has three important functions.

First, it provides a bounded and scale-controlled comparison between pocket and ligand summaries. This avoids treating heterogeneous descriptors as directly comparable before normalization.

Second, it transforms ligand–pocket comparison into a kernel-geometric object, consistent with kernel learning, Gaussian-process similarity, and positive-definite matrix theory [33,35,36].

Third, it supplies the weighted graph on which MQWalk topology validation is performed. In this way, MQWalk is not an independent non-functional module. It is mathematically downstream of the same operator geometry that stabilizes the rank frontier.

The operator-kernel layer is shown explicitly in Figure 2, where pocket and ligand channels are mapped to PSD operators and compared through trace overlap. Figure 6 further shows how this layer supports the construction of an operator-overlap graph for topology validation. The Hyper Mega Core topology map in Figure 10 also belongs to this mathematical layer: its nodes represent pharmacophore-role modules and descriptor-space compatibility relations, not atom-exact scaffold fusions. Thin teleport/refinement edges in Figure 10 indicate kernel-space mapping for ranking and clustering rather than synthetic merging.

This distinction is important because it reduces the risk of conceptual overstatement. The HMC is not a proposed chemical structure. It is a numerical consensus geometry used for docking-oriented pattern recognition, kernel mapping, and audit-forward re-ranking.

2.1.9.3. Robust Conditioning and Avogadro-Anchored Numerical Damping

The third innovation is the use of a shared conditioning layer before branch fusion. Because docking, DFT descriptors, QMC energies, graph-propagation values, and residue/contact fingerprints have different numerical scales, direct summation would be unstable. GLYBATOMAQ™ therefore applies robust standardization,

$$z_j = (x_j - \mathrm{med}_j)/(\mathrm{MAD}_j + \varepsilon) \tag{42}$$

followed by deterministic Avogadro-anchored damping,

$$\tilde{x_j} = NA^{-}\alpha\ z_j \tag{43}$$

and bounded transformation,

$$\hat{x_j} = \tanh(\tilde{x_j}) \tag{44}$$

Here N_A is the CODATA Avogadro constant [43]. It is used only as a reproducible numerical scale anchor, not as a physical claim about particle-counting in the ranking process. The damping step prevents unstable descriptor channels from dominating the fused score, while the hyperbolic tangent bounds extreme values. This is particularly important in shortlist settings, where a small number of candidates can produce heavy-tailed descriptor distributions.

This conditioning stage is shown in Figure 2 and Figure 5B. Its role is methodological rather than cosmetic: before DFT, QMC, and MQWalk evidence can be fused, each feature channel must be made robust, bounded, and reproducible.

2.1.9.4. Integrated DFT Refinement as Deterministic Electronic-Structure Evidence

The DFT branch contributes deterministic electronic-structure evidence. It is grounded in the Hohenberg–Kohn and Kohn–Sham formalism [44,45], with interpretation following standard DFT treatments [46,47]. For a logged pocket–ligand geometry $\chi^{\star}(P, L)$, the Kohn–Sham equation is written as

$$\hat{H}KS\ \psi_i = \varepsilon_i\ \psi_i \tag{45}$$

with

$$\varrho(r) = \Sigma_i\ f_i\ |\psi_i(r)|^2 \tag{46}$$

and the variational stationarity condition

$$\delta E[\varrho] - \mu(\int \varrho\ dr - N)/\delta\varrho = 0 \tag{47}$$

The DFT branch does not assert that the computed electronic energy is a complete binding free energy. Instead, it extracts a deterministic descriptor vector,

$$\Phi DFT = (E, \Delta HL, \mu, \eta, \omega, q, \mu^{d}_{ip}, cSCF) \tag{48}$$

where E_{DFT} is an electronic-energy proxy, Δ_{HL} is a frontier-orbital gap or related descriptor, μ may represent an electronic chemical potential proxy, η a hardness-like descriptor, ω an electrophilicity-like descriptor, q_{stats} charge-distribution summaries, $\Delta\rho_{\mathrm{contact}}$ local density/contact descriptors, and $\mathrm{SCFdiag}$ convergence diagnostics.

The DFT branch score is

$$SDFT = \beta_0 + \Sigma_j\ \beta_j\ \Phi DFT{,}j \tag{49}$$

Its reliability gate is

$$rDFT = \Pi_m\ rDFT{,}m \tag{50}$$

A typical decomposition is

$$rDFT = rSCF\ r_{geom}\ r^{d}_{omain}\ r_{complete} \tag{51}$$

where r_{SCF} records self-consistent-field convergence, r_{geom} records geometry stability, r_{domain} records whether the descriptor vector lies within the calibrated shortlist domain, and r_{desc} records descriptor completeness. If any component is weak, the DFT contribution is down-weighted.

This is an important mathematical-chemistry safeguard. A ligand should not be promoted simply because a DFT descriptor appears favorable if the underlying calculation is unstable. Figure 5A visualizes DFT-derived descriptor distributions and descriptor-plane separation, while Figure 5B shows how the DFT branch enters the fusion logic. In this sense, the DFT branch is a deterministic near-tie discriminator rather than a universal rescoring function.

2.1.9.5. Integrated QMC Refinement as Stochastic Energetic Auditing

The QMC branch contributes stochastic energetic auditing with explicit uncertainty. It follows the general QMC logic described by Foulkes et al. [48] and Needs et al. [49]. The starting variational energy expression is

$$EVMC[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle\ /\ \langle \Psi | \Psi \rangle \tag{52}$$

A diffusion-like refinement view may be written as

$$-\partial\Psi(R,\tau)/\partial\tau = (\hat{H} - ET)\Psi(R,\tau) \tag{53}$$

where the imaginary-time propagation interpretation is used as a computational route toward lower-energy wavefunction components. In GLYBATOMAQ™, the purpose is not full screening-scale QMC for all ligands. Instead, QMC is applied selectively where rank movement is sensitive, where DFT separation is insufficient, or where topology diagnostics suggest that a candidate's promotion requires additional energetic scrutiny.

The QMC branch returns

$$\Phi QMC = (E\_QMC, \sigma QMC, CI\_QMC, M_{eff}, Var(EL)) \quad (54)$$

where $\hat{E}_{\mathrm{QMC}}$ is the estimated energy, σ_{QMC} is the stochastic uncertainty estimate, CIwidth is a confidence-interval width, ESS is an effective sample-size diagnostic, and VarLocal is a local-energy variance diagnostic.

The conservative QMC score is

$$SQMC = -E\_QMC - \lambda\sigma\ \sigma QMC - \lambda CI\ CI\_QMC \quad (55)$$

The minus sign before $\hat{E}_{\mathrm{QMC}}$ makes more favorable lower-energy estimates contribute positively after score alignment, while the two penalty terms prevent uncertain estimates from being overrewarded. This is the core uncertainty-aware feature of the QMC branch.

A reliability gate is also applied,

$$rQMC = r_{sample}\ r_{var}\ rCI\ r_{conv} \quad (56)$$

with a possible factorization,

$$rQMC = I(M_{eff} \geq M_{min})\ I(\sigma \leq \sigma_{max})\ I(CI \leq CI_{max})\ r_{conv} \quad (57)$$

Here r_{ESS} measures sampling adequacy, r_{var} penalizes inflated local-energy variance, r_{CI} penalizes wide intervals, and r_{conv} records convergence stability. A candidate can therefore receive a favorable expected QMC value but still fail to receive strong rank support if the uncertainty is too large.

This branch is summarized in Figure 1A, Figure 2, Figure 3, and Figure 5B. The mathematical-chemistry innovation is not that QMC is used as a brute-force replacement for docking. It is that stochastic electronic uncertainty is explicitly inserted into rank movement. This formally resolves a common weakness of virtual screening pipelines: candidates are often ordered as if their scores were exact, even when the score-generating procedure has substantial uncertainty.

2.1.9.6. DFT/QMC Complementarity and Nonredundancy

DFT and QMC are not redundant in GLYBATOMAQ™. They answer different rank-audit questions.

The DFT branch asks:

$$QDFT(L_a, L_\beta): SDFT(L_a) - SDFT(L_\beta) > \tau D \quad (58)$$

The QMC branch asks:

$$QQMC(L_a, L_\beta): SQMC(L_a) - SQMC(L_\beta) > \tau Q \quad (59)$$

This distinction can be written as a two-stage refinement criterion. A ligand L_a may be preferred over L_b after DFT if

$$L_a \succ DFT\ L_\beta \Leftrightarrow SDFT(L_a) - SDFT(L_\beta) \geq \delta DFT \quad (60)$$

but this preference is only strongly supported if the uncertainty-penalized QMC comparison also satisfies

$$L_a \succ QMC\ L_\beta \Leftrightarrow SQMC(L_a) - SQMC(L_\beta) \geq \delta QMC \quad (61)$$

where τ_{QMC} is a predeclared uncertainty-aware separation threshold. More explicitly, using the conservative QMC score,

$$[-\hat{E}_a - \lambda\sigma_a - \lambda CI\ CI_a] - [-\hat{E}_\beta - \lambda\sigma_\beta - \lambda CI\ CI_\beta] \geq \delta QMC \quad (62)$$

This makes the refinement decision explicitly uncertainty-aware. A ligand with a favorable mean estimate but a large confidence width may fail to outrank a slightly less favorable but more stable candidate. This behavior is essential for responsible mathematical-chemistry ranking.

Figure 5A shows how DFT descriptor distributions can expose near-tie separation in descriptor space. Figure 5B shows how DFT and QMC are not merged prematurely but enter the reliability-aware fusion gate as distinct channels. Figure 6 extends this logic by linking branch-level scores to hotspot landscapes, topology validation, and residue-resolved energetic fingerprints.

2.1.9.7. MQWalk Topology Validation as a Mathematical Consistency Check

The fourth innovation is MQWalk topology validation. MQWalk is not introduced as a physical quantum-biological mechanism. It is a graph-propagation diagnostic on the normalized operator-overlap graph, consistent with the caution required when borrowing quantum terminology in biological contexts [50-52].

From the normalized kernel $\bar{K}(P,L)$, GLYBATOMAQ™ constructs a bipartite adjacency matrix,

$$A = [[0, \tilde{K}], [\tilde{K}^T, 0]] \tag{63}$$

The degree-normalized adjacency is

$$AN = D^{-1}/2\ A\ D^{-1}/2 \tag{64}$$

where D is the diagonal degree matrix. To avoid unstable behavior from disconnected or weakly connected components, teleport-style stabilization is applied,

$$A\gamma = (1\text{-}\gamma)AN + \gamma\ 11^T/|V| \tag{65}$$

The Avogadro-conditioned Laplacian is then

$$\tilde{L} = NA^{-}\alpha(I - A\gamma) \tag{66}$$

The continuous-time propagation operator is

$$U(t) = \exp(-it\tilde{L}) \tag{67}$$

and the resulting node mass is

$$p_j(_t) = |\langle j|U(t)|\psi_0\rangle|^2 \tag{68}$$

These equations define an algorithmic propagation process on the ligand–pocket overlap graph. They do not assert that the ligand or protein undergoes coherent quantum transport. The term MQWalk is therefore used in a mathematical-computational sense: it is a continuous-time quantum-walk-style propagation diagnostic applied to a PSD-kernel graph.

The MQWalk branch tests whether a candidate's promotion is topologically plausible. If DFT and QMC favor a ligand but the ligand is isolated in the operator-overlap graph, or if its propagated mass disperses away from the relevant pocket neighborhood, the promotion is penalized. This is formalized as

$$SMQW = \lambda_p\ s_{persist} - \lambda\Delta\ \Delta_{topo} \tag{69}$$

Here MQWscore (P,L) summarizes walk-supported proximity or persistence, and $\Delta_{topo}(P,L)$ penalizes disagreement between branch-level promotion and graph-level support.

A possible topology-discordance term is

$$\Delta_{topo} = \| z_{\beta ranch} - z_{9raph} \|_2 \tag{70}$$

or, in a normalized form,

$$\Delta_{topo} = 1 - \mathrm{corr}(z_{\beta ranch}, z_{9raph}) \tag{71}$$

where $\mathbf{s}_{branch}$ is the vector of DFT/QMC-supported candidate scores and $\mathbf{s}_{walk}$ is the vector of topology-supported scores. Either form makes the same point: topology is used to audit whether refinement-supported promotions are coherent with the candidate's neighborhood in operator space.

MQWalk diagnostics are shown in Figure 1B–C, Figure 2, Figure 3, and Figure 6. The heatmaps, transport traces, bin summaries, and topology snapshots are not presented as physical observables. They are inspection surfaces for rank stability, dispersion, concentration, and neighborhood consistency.

2.1.9.8. Integrated DFT/QMC/MQWalk Fusion

The final fused score is written as

$$S^{+} = w_0 S\theta + wD_r DSD + wQ_r QSQ + wMSM - \Pi\_diag \qquad (72)$$

This equation is the principal mathematical integration point of the GLYBATOMAQ™ methodology. Each term has a distinct function.

The term $w_0 S_\theta(P, L)$ preserves the upstream pattern-recognition and docking-oriented ranking signal. This prevents the refinement layer from discarding the original screening context.

The term

$$\Delta SDFT = wD\ rDFT\ SDFT \qquad (73)$$

adds deterministic electronic-structure evidence only when DFT reliability is adequate.

The term

$$\Delta SQMC = wQ\ rQMC\ SQMC \qquad (74)$$

adds stochastic energetic support but penalizes uncertainty and low sampling reliability.

The term

$$\Delta SMQW = wM\ SMQW \qquad (75)$$

adds topology-consistency support from the operator-overlap graph.

The penalty

$$\Pi\_diag = \Sigma_h\ \lambda_h\ I_h \qquad (76)$$

prevents unreliable or contradictory evidence from being hidden inside the final score.

A possible expanded penalty form is

$$\Pi\_diag = \lambda D({}_1 - {}_r DFT) + \lambda Q({}_1 - {}_r QMC) + \lambda T\Delta_{topo} + \lambda P\ I_{prov} \qquad (77)$$

This penalty structure is important because it makes uncertainty, convergence failure, topology discordance, and missing provenance part of the decision rule. A candidate cannot be promoted silently if the evidence supporting that promotion is incomplete or unstable.

The fusion logic is shown in Figure 2, Figure 5B, and Figure 6. Figure 5B is especially important because it presents the sequence

Scale → Expand → Aggregate → Score → Compare → Refine → Validate → Re-rank (78)

This sequence clarifies why the method is not an arbitrary stack of computational tools. Each stage transforms the ranking object in a controlled way, and each transformation is auditable.

2.1.9.9. Residue-Resolved Energetic Fingerprints as an Interpretability Bridge

A further mathematical-chemistry feature is the use of residue/contact decomposition as an interpretability bridge between scalar scores and structural evidence. For ligand L, the interface energy fingerprint may be written as

$$E_{int}(L) = \Sigma_r \in \boldsymbol{\mathcal{V}}\ e_r(L) \qquad (79)$$

where R is the residue/contact vocabulary and $\Delta E_r(L)$ is the residue-resolved energetic contribution.

The corresponding sparse vector is

$$e(L) = (e_{r1}(L), \ldots, e_{rm}(L)) \qquad (80)$$

The nonzero support,

$$supp(eL) = r \in \boldsymbol{\mathcal{V}} : |e_r(L)| > \tau_e \qquad (81)$$

defines the ligand's hotspot signature. Two ligands may have similar total interface scores but different hotspot supports. Conversely, two ligands may share a hotspot signature but differ in the magnitude or balance of stabilizing and destabilizing terms.

This is shown in Figure 6. For example, model1_2gli–GLYBATOMAQ-HSX-cons1 is reported with a total energy of approximately -100.1 and strong stabilizing terms including $\Delta E_{\mathrm{V-M-GLN154}} = -13.8519$, $\Delta E_{\mathrm{V-S-ARG163}} = -10.0914$, $\Delta E_{\mathrm{V-M-GLU11}} = -9.75054$, and $\Delta E_{\mathrm{V-S-TRP10}} = -8.23134$, along with a counteracting positive term $\Delta E_{\mathrm{V-S-HIS140}} = +6.76114$. In the same figure, model1_2gli–GLYBATOMAQ-HSX-cons2 is reported with a total energy of approximately -97.5, with a different stabilizing pattern dominated by $\Delta E_{\mathrm{V-M-HIS164}} = -14.0114$ and $\Delta E_{\mathrm{V-M-SER226}} = -11.1797$, together with other stabilizing and destabilizing terms.

The interpretive value is that rank behavior can be related to energetic wiring. If two candidates are close in total score but have different residue supports, their rank comparison is chemically nontrivial. The residue fingerprint therefore complements DFT/QMC/MQWalk refinement by giving a structural explanation for why a candidate is supported, weakly supported, or discordant.

The same logic is compatible with the interaction-envelope persistence statistic shown in Figure 3, Figure 4, and Figure 6,

$$\pi_{c},s,t(L) = N_{c},s,t(L)/(N_{s},t+\varepsilon) \quad (82)$$

where c denotes an interaction class, S receptor or structural states, T time or sampling indices, and $I_{c,s,t}^{(k)}$ an observed interaction event. This statistic distinguishes recurrent interaction patterns from single-pose artifacts.

2.1.9.10. Consensus Leaderboards and Pocket-Dependent Reversals

Because GLIPR1 pockets may induce different rankings, GLYBATOMAQ™ reports pocket-specific and consensus leaderboards. For a pocket set $\mathcal{P}^{(T)}$, the consensus score is

$$S_{cons}(L) = \mathrm{Agg}P\in\mathcal{P}\ S^{+}(P,L) \quad (83)$$

where Agg is a predeclared aggregation rule. Possible choices include a mean, trimmed mean, median, rank product, or Borda-type rule. The important requirement is that the rule is fixed before final interpretation.

A rank-product version may be written as

$$SRP(L) = -\Sigma P\in\mathcal{P}\ \log(R^{+}(P,L)+\varepsilon) \quad (84)$$

where lower ranks are better and the negative sign aligns the value with a larger-is-better scoring convention. A Borda-style version may be written as

$$SB_{or}{}^{d}{}_{a}(L) = \Sigma P\in\mathcal{P}(|\mathcal{L}P| - R^{+}(P,L)) \quad (85)$$

where N_L is the number of ligands. These consensus rules make pocket dependence explicit rather than hiding it.

Figure 8 illustrates why this matters. The AutoDock Vina matrix shows pocket-by-pocket variation, including cases where leading candidates achieve their best values in different pockets. Figure 9 similarly shows that DockThor affinity and component energies can differ across ligands even when affinity values are close. For example, the DockThor panel reports GLYBATOMAQ-HSX-cons2 as the strongest DockThor affinity candidate in that set, while other candidates show distinct energy-component patterns. In GLYBATOMAQ™, such differences are not treated as contradictions. They are treated as evidence that rank interpretation must remain pocket-aware and component-aware.

2.1.9.11. Relationship to the HMC and HSX Consensus Representations

The Hyper Mega Core topology map in Figure 10 and the HSX consensus panels in Figure 3, Figure 4, and Figure 5 are interpreted as numerical and chemistry-safe reference structures for ranking, clustering, and re-ranking. They are not atom-exact mega-scaffolds and not synthesis prescriptions. Figure 10 extends this HMC interpretation by making the constructive two-dimensional assembly certificate explicit: admissible HSX candidates are paths in a role-layered graph whose planar geometries remain constrained by distance, angle, non-overlap, and kernel-compatibility conditions.

In Figure 10, nodes represent pharmacophore-role modules,

$$VHMC = M_{rec}, M_{link}, M_{int}, M_{space}, M_{sol}, M_{cap} \quad (86)$$

Edges represent descriptor-space compatibility. Thin teleport/refinement links connect comparator anchors to the HMC to show kernel-space mapping. This is directly analogous to the MQWalk teleport mechanism, but at the representation level rather than the final validation graph.

The HSX-cons1-4 structures shown in Figure 3 and Figure 5 are fixed reference anchors. Their role is to support consistent comparison, not to imply that all pharmacophore modules are chemically mergeable. The HMC/HSX layer is therefore a controlled reference geometry for rank-centric pattern recognition, kernel mapping, and audit-forward re-ranking.

2.1.9.12. Relationship to External Docking and Network-Pharmacology Examples

Figures 1 and 2 include external computational examples to contextualize the rank-first philosophy. In Figure 1, the MMP3 inhibitor example is used to illustrate that docking, molecular dynamics, and experimental dose comparisons are often interpreted in practice as rank-bearing evidence rather than as absolute equivalences. The reported ordering of ilomastat derivatives, MD-refined ΔG_{bind} values, and inhibition comparisons at fixed concentrations are used as conceptual anchors for why near-tie regions require careful discrimination.

In Figure 2, the vortioxetine glioblastoma network-pharmacology example similarly illustrates rank-centric interpretation. Target harvesting, PPI reduction, enrichment, and docking values produce a structured prioritization pipeline. In the GLYBATOMAQ™ framing, those values are not imported as proof of efficacy. They are used to show how modern computational drug-repurposing pipelines often produce ranked hypotheses requiring downstream audit.

The methodological relevance is that GLYBATOMAQ™ provides an explicit mathematical mechanism for such auditing. The DFT branch asks whether local electronic descriptors support a candidate; the QMC branch asks whether energetic support survives uncertainty; the MQWalk branch asks whether the candidate remains coherent in the operator-overlap graph; and the fusion layer asks whether the final rank movement is reliable.

2.1.9.13. Why the Complexity is Necessary and Controlled

The mathematical-chemistry usefulness of GLYBATOMAQ™ depends on whether each computational layer contributes a distinct, testable function. The framework is organized so that every layer has an explicit role in the rank-audit chain.

The docking layer is necessary because it supplies the fixed-protocol structural frontier:

- The PSD operator-kernel layer is necessary because it converts heterogeneous pocket and ligand descriptors into a controlled similarity geometry.
- The DFT branch is necessary because near-tie candidates may differ in local electronic structure even when docking scores are similar.
- The QMC branch is necessary because deterministic descriptors alone do not quantify stochastic energetic uncertainty.
- The MQWalk branch is necessary because a favorable branch score should not be allowed to create a rank promotion that is inconsistent with the candidate's learned ligand–pocket neighborhood.
- The diagnostic penalty layer is necessary because unreliable calculations should not silently influence the final leaderboard.

Thus, the method is not complicated because it adds unrelated modules. It is layered because rank movement in molecular screening has multiple failure modes:

$$\boldsymbol{\mathcal{F}}\oint_{\text{ail}} = \text{near-tie ambiguity, descriptor drift, high QMC variance, topology isolation,} \\ \text{missing provenance} \tag{87}$$

GLYBATOMAQ™ addresses these failure modes through explicit branch diagnostics.

2.1.9.14. Methodological Limits

Despite its mathematical structure, GLYBATOMAQ™ has explicit limits. These limits are part of the methodology and should not be treated as weaknesses hidden from interpretation.

- First, GLYBATOMAQ™ does not replace experimental binding assays. The final leaderboard is a prioritization hypothesis, not a measurement of affinity, efficacy, toxicity, or clinical performance.
- Second, docking scores from AutoDock Vina, AutoDock4, Glide, DockThor-VS, or related tools are not interpreted as experimental binding free energies [6-16]. They define a protocol-specific ranking frontier.
- Third, DFT descriptors do not by themselves constitute a complete protein–ligand binding thermodynamic model. They provide deterministic local electronic-structure evidence under logged geometries [44-47].
- Fourth, QMC is used selectively as an uncertainty-aware audit layer. It is not claimed to provide exhaustive screening-scale electronic refinement for all candidates [48,49].
- Fifth, MQWalk is an algorithmic topology diagnostic. It is not a claim of physical quantum transport in GLIPR1 or in ligand recognition. The quantum-walk terminology is used by analogy to continuous-time propagation on a graph, consistent with the need for caution in quantum-biological interpretation [50-52].
- Sixth, the HMC and HSX consensus representations are chemistry-safe reference geometries for comparison, ranking, and audit. They are not synthesis prescriptions and not validated therapeutic structures.
- Seventh, the method is dependent on the quality and consistency of upstream inputs. If pocket definitions, ligand conformers, docking poses, or descriptor provenance are poor, the refinement layer cannot convert poor inputs into validated biological conclusions. It can only expose instability, discordance, or weak support.

2.1.9.15. Summary of the Mathematical-Chemistry Contribution

The final methodological contribution can be summarized as the following rank-audit map:

$$\boldsymbol{\mathcal{A}}(\mathrm{P,L}) = \mathrm{R}_0,\ \bar{\mathrm{K}},\ \Phi\mathrm{DFT},\ \Phi\mathrm{QMC},\ \mathrm{SMQW},\ \Pi\_\mathrm{diag},\ \mathrm{S}^{+},\ \Delta\mathrm{R} \tag{88}$$

This object is the practical output of GLYBATOMAQ™. It records the initial rank, the operator-kernel geometry, the electronic refinement evidence, the uncertainty-aware energetic audit, the topology-validation score, the diagnostic penalties, and the final rank shift.

In this formulation, a candidate is not promoted because one number is favorable. A candidate is promoted only when the promotion is supported by a chain of evidence: The decision rule is explicitly encoded by Eq. (250) and by the reporting definitions in Table 12 and Figure 13.

$$\mathrm{Promote(L)} \Leftrightarrow \Delta\mathrm{R}{>}0 \wedge \mathrm{rDFT,rQMC} \geq \tau_{\mathrm{r}} \wedge \Delta_{\mathrm{topo}} \leq \tau_{\mathrm{topo}} \tag{89}$$

This is the central research-level mathematical-chemistry innovation of the GLYBATOMAQ™ methodology. It reframes GLYBATOMAQ™ as an audit-forward ranking formalism for GLIPR1 ligand prioritization, with explicit equations, reliability gates, uncertainty penalties, graph-topology validation, and figure-linked interpretability across Figure 10, Figure 1, Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, and Figure 9.

The method therefore avoids conceptual inflation while preserving the useful mathematical content. It does not claim that the quantum-enriched refinement circuit proves binding or efficacy. It claims that, within a fixed computational protocol, rank movement can be made more transparent, more reproducible, more uncertainty-aware, and more topologically constrained. That is the intended mathematical-chemistry usefulness of GLYBATOMAQ™. This interpretation is consistent with the visual and tabular reporting layer introduced in Figures 10-14 and Tables 12-15.

Algorithm 1. GLYBATOMAQ™ DFT/QMC/MQWalk refinement and reliability-aware re-ranking — HSX-cons1–4 IUPAC strings (SI Appendix II Pseudocode)

Input

- ***Pocket set*** $\mathcal{P} = \{P\}$ *(protocol-logged GLIPR1 pockets).*
- ***Candidate sets*** $\{\mathcal{L}_P\}$ *with ligands* L *carrying:*
    - ○ *base score* $S_\theta(P, L)$*,*
    - ○ *docking oracle* $D(P, L)$ *(logged; optional in fusion),*

  - *PSD operators* $\rho_P, \rho_L \geq 0$,
  - *normalized overlap*

$$M_{90} = \mathrm{Log}_{90}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{90}$$

  - $Conf(L)$ *and* $pose(P, L)$ *(when available).*
- ***Shortlist policy*** Π_{short}.
- ***Protocols/configs**:* $\Omega_{DFT}, \Omega_{QMC}, \Omega_{MQW} = \{\alpha, \gamma, t, \lambda_{walk}, \lambda_{disc}, \dots\}$, *fusion* $\Omega_{FUSE} = \{w_0, w_{DFT}, w_{QMC}, w_{MQW}, \Pi_{diag}\}$, *reliability* Ω_{REL}, *logging* Ω_{LOG}.

Output

- *Pocket-wise leaderboard(s) ranked by* $S_{final}^{(\;+\;)}(P, L)$.
- *Optional consensus leaderboard(s)* $S_{cons}^{(\;+\;)}(L)$.
- *Rank-shift diagnostics* $\Delta R(P, L) = R_0(P, L) - R_+(P, L)$.
- *Branch contribution tables* $\Delta S_{DFT}, \Delta S_{QMC}, \Delta S_{MQW}$, *topology disagreement* Δ_{topo}, *flags, and run manifest* M.
- ***HSX-cons1234 Mega-Core dashboard export** (when those candidates appear in the processed shortlist/union) including a **named output appendix**:*

O-HSX

- ***GLYBATOMAQ-HSX-cons2 — Preferred IUPAC Name** = 2-[(4-amino-1,2,4-triazinan-1-yl)amino]-N-[2-({8-carbamoyl-4-oxo-octahydroimidazolidino[4,3-d][1,2,3,5]tetrazin-3-yl}methyl)-5-oxa-1-azabicyclo[2.1.0]pentan-4-yl]-1-[3-({1-[3′′-formamido-12′-(2-methylpropyl)-2,2′,8′-trioxo-octahydro-1′′λ^{4},5′λ^{5}-trispiro[indole-3,11′-[1,7]diaza-[5λ^{5}]phosphatricyclo[7.3.0.$0^{3,7}$]dodecane-5′,4′′-[1,2,4]thiazaphosphinane-1′′,1′′′-[1,2,4]thiadiazolidine]-3′′′-yl]ethyl}carbamoyl)-1-(methylamino)propyl]-5-(2-hydrazinylethyl)piperidine-4-carboxamide*
- ***GLYBATOMAQ-HSX-cons1 — Preferred IUPAC Name** = 6-{2-[(4-amino-1,2,4-triazinan-1-yl)amino]-4-{[2-({8-carbamoyl-4-oxo-octahydroimidazolidino[4,3-d][1,2,3,5]tetrazin-3-yl}methyl)-5-oxa-1-azabicyclo[2.1.0]pentan-4-yl]carbamoyl}-5-(2-hydrazinylethyl)piperidin-1-yl}-5-chloro-N-{1-[6′′-chloro-3′′-formamido-12′-(2-methylpropyl)-2,2′,8′-trioxo-octahydro-1′′λ^{4},5′λ^{5}-trispiro[indole-3,11′-[1,7]diaza-[5λ^{5}]phosphatricyclo[7.3.0.$0^{3,7}$]dodecane-5′,4′′-[1,2,4]thiazaphosphinane-1′′,1′′′-[1,2,4]thiadiazolidine]-3′′′-yl]ethyl}-1,3-diazinane-4-carboxamide*
- ***GLYBATOMAQ-HSX-cons4 — Preferred IUPAC Name** = 1-{2-[(4R,6S)-4-{[(2S)-4-[(4-amino-decahydropteridin-2-yl)({8-carbamoyl-4-oxo-octahydroimidazolidino[4,3-d][1,2,3,5]tetrazin-3-yl})amino]azetidin-2-yl]carbamoyl}-1-[5-chloro-6-({1-[(3′′R,11′′R,12′′S)-6′,12′′-dichloro-3′-formamido-2′′,8′′-dioxo-1λ^{4},4′λ^{5}-trispiro[1,2,4-thiadiazolidine-1,1′-[1,2,4]thiazaphosphinane-4′,5′′-[1,7]diaza-[5λ^{5}]phosphatricyclo[7.3.0.$0^{3,7}$]dodecane-11′′,4′′′-[12]oxa-[3,5]diazatricyclo[7.4.0.$0^{2,6}$]tridecane]-3-yl]ethyl}carbamoyl)-1,3-diazinan-4-yl]-6-[(1,2,4-triazinan-1-yl)amino]piperidin-3-yl]ethyl}hydrazinium*
- ***GLYBATOMAQ-HSX-cons3 — Preferred IUPAC Name** = 6-{2-[(4-amino-1,2,4-triazinan-1-yl)amino]-4-{[2-({8-carbamoyl-4-oxo-octahydroimidazolidino[4,3-d][1,2,3,5]tetrazin-3-yl}methyl)-5-oxa-1-azabicyclo[2.1.0]pentan-4-yl]carbamoyl}-5-(2-hydrazinylethyl)piperidin-1-yl}-N-{1-[7-({amino[2-(3-{[10-(aminomethyl)-dodecahydropiperazino[1,2-a][1,3,5,7]tetrazonin-9-yl]amino}aziridin-2-yl)hydrazin-1-yl]methyl}amino)-10-chloro-9-(trifluoromethyl)-5λ^{4}-thia-1,3,6-triazaspiro[4.5]decan-2-yl]ethyl}-5-chloro-1,3-diazinane-4-carboxamide*

Procedure—Pseudocode (SI Appendix II)

Algorithm 1. GLYBATOMAQ™ quantum-enriched re-ranking and audit workflow

Purpose and mathematical-chemistry rationale

Algorithm 1 formalizes GLYBATOMAQ™ as a **rank-centric, audit-forward mathematical-chemistry workflow**. Its purpose is not to replace docking, experimental binding assays, molecular dynamics, free-energy perturbation, or prospective validation. Instead, it provides a controlled post-docking re-ranking layer in which candidate promotions and demotions are made traceable to four explicitly separated sources of evidence:

$$M_{91} = \mathrm{Log}_{91}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{91}$$

This formalization keeps the architecture compact and function-driven. Each component is assigned a specific mathematical function. Docking provides the fixed-protocol upstream ranking frontier; positive-semidefinite operator kernels provide a stable ligand-pocket comparison geometry; DFT contributes deterministic electronic-structure descriptors; QMC contributes uncertainty-aware energetic auditing; MQWalk tests whether rank promotions remain consistent with the learned topology of the ligand-pocket similarity graph; and the final fusion rule converts these signals into a reliability-gated rank movement rather than an unqualified score.

The workflow is therefore best interpreted as a **mathematical audit of rank displacement**. The central question is not simply whether ligand L receives a high final score. The central question is whether the transition

$$M_{92} = \mathrm{Log}_{92}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{92}$$

is supported by stable branch evidence, controlled uncertainty, and topology-consistent neighborhood behavior. This is the methodological position illustrated in SI Appendix I Figure 1, SI Appendix I Figure 2, SI Appendix I Figure 5, and SI Appendix I Figure 6.

Consistent with benchmark-aware docking practice, docking outputs are treated as comparative ranking cues rather than literal affinity estimates [9-16]. The rank-centric framing is also compatible with recent Pattern Recognition work on ranking-based learning, local-kernel graph learning, informative graph selection, dynamic graph representation, adaptive graph propagation, and molecular graph contrastive learning [56-61]. Algorithm 1 therefore converts a conventional docking shortlist into an auditable mathematical-chemistry object.

Algorithm 1. GLYBATOMAQ™ quantum-enriched re-ranking and audit workflow

Input.

A protocol-logged pocket set

$$M_{93} = \mathrm{Log}_{93}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{93}$$

pocket-specific ligand sets

$$M_{94} = \mathrm{Log}_{94}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{94}$$

a base GLYBATOMAQ score $S_\theta(P,L)$, docking outputs $D(P,L)$, operator descriptors, conformer and pose provenance, and protocol bundles for DFT, QMC, MQWalk, fusion, reliability scoring, shortlisting, and reporting.

The input may include protein structures from the Protein Data Bank [1], GLIPR1 structural context from Asojo et al. [2], GLIPR1-related structures 3Q2U [3] and 3Q2R [4], pocket definitions supported by Fpocket [5], docking outputs from AutoDock Vina [6], AutoDock4 [7], Glide [8], DockThor-VS [13], or equivalent fixed-protocol screening outputs. Docking scores are treated as rank-bearing computational signals, not experimental affinity values, consistent with the docking literature [9-16].

Output.

Pocket-wise leaderboards, consensus leaderboards, explicit rank-shift diagnostics, branch-level contributions, DFT reliability flags, QMC uncertainty summaries, MQWalk topology flags, optional HSX dashboard outputs, provenance appendices, and a complete run manifest

$$M_{95} = \mathrm{Log}_{95}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{95}$$

The primary output is not a single score but an audit bundle,

$$M_{96} = \mathrm{Log}_{96}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{96}$$

This audit bundle is the main mathematical-chemistry object produced by Algorithm 1.

Step 0. Initialize manifest, define scope, and lock protocol settings

Initialize the run manifest

$$M_{97} = \mathrm{Log}_{97}(P,L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{97}$$

The manifest records all settings needed to replay the ranking decision. These include receptor structures, pocket definitions, ligand identifiers, conformer-generation settings, docking software and parameters, scoring conventions, descriptor-generation rules, kernel construction rules, DFT settings, QMC settings, MQWalk parameters, reliability thresholds, fusion weights, and reporting options.

The manifest is not a bookkeeping afterthought. It is required because the scientific output of GLYBATOMAQ™ is a **rank movement with provenance**. A ligand promotion cannot be evaluated if the protocol that produced it is not replayable. This design follows benchmark-oriented virtual-screening practice, where docking and scoring functions must be interpreted in relation to fixed protocols and controlled test sets [9,10,14-16].

At this stage, the algorithm also locks the interpretation boundary:

$$M_{98} = \mathrm{Log}_{98}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{98}$$

meaning that docking output is not treated as an experimental binding free energy. Instead,

$$M_{99} = \mathrm{Log}_{99}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{99}$$

This statement is essential for avoiding overclaiming. The algorithm uses docking to define a frontier, not to assert biochemical efficacy.

The initialization stage corresponds to the left side of the workflow in SI Appendix I Figure 1A and SI Appendix I Figure 2, where upstream docking/oracle scoring is separated from downstream operator-kernel stabilization and refinement.

Step 1. Construct pocket-wise shortlists and pre-refinement ranks

For each pocket $P \in \mathcal{P}$, compute or retrieve the base GLYBATOMAQ score

$$M_{100} = \mathrm{Log}_{100}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{100}$$

for every ligand $L \in \mathcal{L}_P$. The pocket-wise shortlist is then defined as

$$M_{101} = \mathrm{Log}_{101}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{101}$$

The pre-refinement rank is assigned as

$$M_{102} = \mathrm{Log}_{102}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{102}$$

This stage establishes the reference leaderboard against which all later movement is measured. It is important that the DFT, QMC, and MQWalk branches are not applied indiscriminately to the entire library. They are invoked only after shortlisting because their role is to refine and audit a chemically plausible frontier, not to serve as high-throughput replacement docking engines.

This shortlisting principle also controls computational cost. If N denotes the full library size and $K \ll N$ denotes the shortlist size, then the expensive refinement branches scale with K, not N. The computational logic is therefore

$$M_{103} = \mathrm{Log}_{103}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{103}$$

This is one of the safeguards against uncontrolled complexity. DFT and QMC are not used as non-functional computations across all candidates; they are reserved for rank-sensitive regions where additional evidence can change interpretation.

Near-tie behavior may also be identified at this stage. For example, using a best-pocket docking summary,

$$M_{104} = \mathrm{Log}_{104}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{104}$$

the distance-to-best statistic is

$$M_{105} = \mathrm{Log}_{105}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{105}$$

Small values of Δ_{dock} indicate candidates whose apparent separation may not be reliable under docking alone. These are precisely the cases for which DFT/QMC refinement and MQWalk topology validation are most useful. The same rank-first logic is visible in SI Appendix I Figure 7 and SI Appendix I Figure 8, where pocket-resolved docking values and near-tie gaps are treated as ranking artifacts rather than binding-affinity claims.

Step 2. Assemble candidate objects and harmonize chemical representations

For each pocket–ligand pair

$$M_{106} = \mathrm{Log}_{106}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{106}$$

construct a candidate object

$$M_{107} = \mathrm{Log}_{107}(P, L, \Omega DFT, \Omega QMC, \Omega MQW, \Omega FUSE) \tag{107}$$

A minimal candidate object is

$$M_{108} = \mathrm{Log}_{108}(P,L, \Omega\mathrm{DFT}, \Omega\mathrm{QMC}, \Omega\mathrm{MQW}, \Omega\mathrm{FUSE}) \tag{108}$$

Here ID_L is the ligand identity; SMILES_L and $\mathrm{SELFIES}_L$ are optional structure encodings [21,22]; $\mathrm{Conf}(L)$ is the conformer ensemble; $\chi^\star(P,L)$ is the selected or harmonized pocket–ligand geometry; $D(P,L)$ is the protocol-fixed docking output; $S_\theta(P,L)$ is the base score; ρ_P and ρ_L are operator summaries; and Π_{prov} stores provenance.

Conformer handling and chemical preprocessing are compatible with MMFF94 [17], RDKit [18], Morgan-style encodings [19], ECFP fingerprints [20], and SMILES-based preprocessing [21]. The representation layer is also compatible with grammar VAE, junction-tree VAE, graph-policy learning, recurrent molecular design, continuous latent molecular design, neural message passing, graph convolution, and attention-based representations [23-32]. These representations are used as admissible encodings, not as independent proof of biological activity.

The geometry harmonization step may be written as

$$\chi^*_{109}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{109}$$

where Θ_{pose} records the pose-selection rule. A common fixed-protocol rule is

$$\chi^*_{110}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{110}$$

Alternative pose-selection rules can be used, but they must be logged in $\mathcal{M}$. This is essential because a rank movement after DFT/QMC refinement depends on the geometry used for electronic evaluation.

The candidate object is therefore not merely a row in a spreadsheet. It is a replayable molecular evidence object. Its role is shown conceptually in SI Appendix I Figure 2, where the LOAD and ENCODE stages initialize a candidate record before kernel mapping and branch refinement.

Step 3. Construct PSD operator summaries and normalized overlap kernels

For each candidate object $O(P,L)$, construct positive-semidefinite operator summaries

$$\chi^*_{111}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{111}$$

A standard construction is

$$\chi^*_{112}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{112}$$

where Z is a feature, descriptor, graph-embedding, hotspot, or interaction matrix. This construction is mathematically useful because it ensures that the resulting operator lies in the positive-semidefinite cone. The PSD structure allows the ligand–pocket comparison to be interpreted through trace overlaps and kernel geometry [33,36].

The unnormalized overlap is

$$\chi^*_{113}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{113}$$

and the normalized overlap is

$$\chi^*_{114}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{114}$$

The normalized kernel $\bar{K}(P,L)$ serves three purposes.

First, it creates a scale-controlled ligand–pocket similarity signal.

Second, it stabilizes the shortlist by comparing candidate objects in an operator space rather than only through scalar docking values.

Third, it defines the weighted graph used for MQWalk topology validation.

This step is the mathematical bridge between pattern recognition and quantum-enriched refinement. It is shown explicitly in SI Appendix I Figure 2, where pocket and ligand channels are mapped to PSD operators, and in SI Appendix I Figure 6, where the operator-overlap graph becomes the basis for topology validation.

To emphasize the mathematical role of this layer, Algorithm 1 may store the following kernel bundle:

$$\chi^*_{115}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi\|\chi-\chi_0\|^2] \tag{115}$$

This kernel bundle is later used by the MQWalk branch and by diagnostic penalties when branch-supported promotions disagree with topology.

Step 4. Apply robust shared conditioning and Avogadro-anchored numerical damping

Because Algorithm 1 integrates heterogeneous evidence channels, all feature values must be conditioned before fusion. For each scalar feature channel f_k, compute the robust standardized value

$$\chi^*_{116}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{116}$$

The median and median absolute deviation are preferred over mean and variance because shortlist distributions are often small, skewed, and sensitive to outliers. This is especially relevant for electronic descriptors, docking components, residue-energy terms, and graph-propagation summaries.

Next, compute the deterministic Avogadro-anchored damping factor

$$\chi^*_{117}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{117}$$

where N_A is the CODATA Avogadro constant [43]. In Algorithm 1, N_A is used only as a reproducible numerical scale anchor. It is not interpreted as a physical molecular-counting term.

The bounded transformed feature is

$$\chi^*_{118}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{118}$$

This transformation prevents isolated extreme values from dominating fusion. The full transformation is therefore

$$\chi^*_{119}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{119}$$

This conditioning step is compatible with the kernel-space and approximation-theoretic foundation of the method [33,36-42]. It is shown in SI Appendix I Figure 2 and SI Appendix I Figure 5B. Its role is not non-functional: without shared conditioning, DFT, QMC, docking, and topology scores would not be numerically comparable.

Step 5. Evaluate the DFT branch as deterministic electronic refinement

For each candidate object $O(P,L)$, run the DFT branch if enabled. The formal basis is the Hohenberg–Kohn and Kohn–Sham framework [44,45], interpreted through standard DFT treatments [46,47].

For the logged geometry $\chi^\star(P,L)$, the Kohn–Sham equation is

$$\chi^*_{120}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{120}$$

with

$$\chi^*_{121}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{121}$$

and the stationarity condition

$$\chi^*_{122}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{122}$$

The DFT branch extracts a descriptor vector

$$\chi^*_{123}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{123}$$

The DFT branch score is

$$\chi^*_{124}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{124}$$

The DFT branch is reliability-gated by

$$\chi^*_{125}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{125}$$

A useful explicit form is

$$\chi^*_{126}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{126}$$

where r_{SCF} records self-consistent-field convergence, r_{geom} records geometry stability, r_{domain} records descriptor-domain validity, and r_{desc} records descriptor completeness.

If the DFT branch is not run, the algorithm returns

$$\chi^*_{127}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{127}$$

If DFT fails, the branch returns

$$\chi^*_{128}(P,L) = \arg\min\chi\,[D(P,L,\chi) + \lambda\chi\lVert\chi-\chi_0\rVert^2] \tag{128}$$

This prevents failed electronic calculations from being silently interpreted as unfavorable chemistry. Failure is treated as diagnostic information, not as a negative score.

The DFT branch contributes deterministic electronic separation in near-tie regions. It does not claim to produce complete protein–ligand thermodynamics. Its diagnostic role is visualized in SI

doi:10.20944/preprints202605.2096.v1

Appendix I Figure 5A through descriptor histograms and scatter diagnostics, and its fusion role is shown in SI Appendix I Figure 5B.

Step 6. Evaluate the QMC branch as uncertainty-aware energetic auditing

For selected candidates, run the QMC branch. The branch follows the uncertainty-aware logic of quantum Monte Carlo methodology [48,49]. The variational energy expression is

$$\chi^*_{129}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{129}$$

A diffusion-like refinement view may be written as

$$\chi^*_{130}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{130}$$

The QMC branch returns

$$\chi^*_{131}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{131}$$

The conservative QMC score is

$$\chi^*_{132}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{132}$$

This equation is central to the algorithm's uncertainty logic. A candidate is not rewarded solely for a favorable estimated energy. It must also have acceptable uncertainty. Large stochastic uncertainty, broad confidence intervals, or poor sampling reduce the QMC contribution.

The QMC reliability is

$$\chi^*_{133}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{133}$$

with a possible factorization

$$\chi^*_{134}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{134}$$

Here r_{ESS} records effective sample size, r_{var} records local-energy variance behavior, r_{CI} records confidence-width acceptability, and r_{conv} records convergence stability.

If QMC is not run, return

$$\chi^*_{135}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{135}$$

If QMC fails or yields unstable uncertainty, return a diagnostic flag such as

$$\chi^*_{136}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{136}$$

The QMC branch is therefore a stochastic audit layer. It is used selectively and conservatively, not as an exhaustive replacement for docking. Its role is summarized in SI Appendix I Figure 1A, SI Appendix I Figure 2, SI Appendix I Figure 3, and SI Appendix I Figure 5B.

Step 7. Evaluate MQWalk topology validation on the operator-overlap graph

Construct the bipartite pocket–ligand graph from the normalized overlap kernel $\bar{K}$. The adjacency matrix is

$$\chi^*_{137}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{137}$$

The degree-normalized adjacency is

$$\chi^*_{138}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{138}$$

where D is the diagonal degree matrix.

To stabilize disconnected or weakly connected components, apply teleport-style mixing,

$$\chi^*_{139}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{139}$$

The Avogadro-conditioned graph Laplacian is

$$\chi^*_{140}(P,L) = \arg\min\chi\ [D(P,L,\chi) + \lambda\chi \|\chi-\chi_0\|^2] \tag{140}$$

The continuous-time walk operator is

$$B_{141}(P,L) = \text{rDFT SDFT + rQMC SQMC + SMQW - }\Pi\text{\_diag} \tag{141}$$

For an initial state $|\psi_0\rangle$, the propagated node mass is

$$B_{142}(P,L) = \text{rDFT SDFT + rQMC SQMC + SMQW - }\Pi\text{\_diag} \tag{142}$$

The MQWalk branch then computes a topology-supported score,

$$B_{143}(P,L) = \text{rDFT SDFT + rQMC SQMC + SMQW - }\Pi\text{\_diag} \tag{143}$$

The topology-disagreement term may be written as

$$B_{144}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{144}$$

or as a normalized disagreement,

$$B_{145}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{145}$$

The first form is rank-displacement based; the second form is vector-agreement based. Both test whether refinement-supported promotions remain consistent with graph topology.

MQWalk is interpreted strictly as an algorithmic topology-validation procedure. Although the notation uses continuous-time quantum-walk formalisms, the algorithm does not claim physical quantum transport in GLIPR1 or ligand recognition. This restricted interpretation is consistent with quantum-information formalism [50] and with broader cautions regarding quantum-biological claims [51,52].

This branch is shown in SI Appendix I Figure 1B–C, SI Appendix I Figure 2, SI Appendix I Figure 3, and SI Appendix I Figure 6. The topology heatmaps, node-mass plots, transport summaries, and bin-resolved diagnostics are inspection surfaces for graph consistency, not physical observables.

Step 8. Compute diagnostic penalties

For each candidate, compute a diagnostic penalty

$$B_{146}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{146}$$

A useful expanded form is

$$B_{147}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{147}$$

The penalty term is a major safeguard against overinterpretation. It prevents unstable branch results from being hidden inside a favorable final score. A candidate may have strong docking and DFT values, but if QMC uncertainty is inflated or MQWalk topology is discordant, that weakness appears explicitly in Π_{diag}.

The diagnostic penalty can also be decomposed for reporting:

$$B_{148}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{148}$$

This decomposition supports transparent leaderboard interpretation. A demotion can be attributed to a specific failure mode, such as DFT instability, QMC uncertainty, topology discordance, or missing provenance.

Step 9. Fuse branch evidence into a reliability-aware final score

Compute the final GLYBATOMAQ™ score as

$$B_{149}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{149}$$

The fusion weights

$$B_{150}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{150}$$

are fixed or logged in $\mathcal{M}$. They must not be tuned post hoc to force a desired leaderboard.

Each term has a distinct function:

$$B_{151}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{151}$$

preserves the upstream rank frontier;

$$B_{152}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{152}$$

adds deterministic electronic-structure evidence only when reliable;

$$B_{153}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{153}$$

adds uncertainty-aware energetic support;

$$B_{154}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{154}$$

adds topology-consistency support;

$$B_{155}(P,L) = rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag \tag{155}$$

removes unsupported or unstable promotions.

The fusion rule therefore makes Algorithm 1 a reliability-aware evidence-integration procedure rather than a simple rescoring layer. Its visual counterpart is shown in SI Appendix I Figure 2 and SI Appendix I Figure 5B.

Step 10. Assign post-refinement ranks and compute rank movement

For each pocket P, assign the post-refinement rank

$$B_{156}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{156}$$

Compute rank movement as

$$B_{157}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{157}$$

Interpretation:

$$B_{158}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{158}$$

The algorithm also computes rank-stability summaries, including

$$B_{159}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{159}$$

and Spearman rank correlation,

$$B_{160}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{160}$$

If the refinement layer dramatically changes the leaderboard, the algorithm requires the movement to be explained through branch evidence. This explanation is recorded as

$$B_{161}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{161}$$

The post-refinement audit is the core output of Algorithm 1. It preserves the manuscript's emphasis on interpretable rank behavior rather than isolated scalar scores, and aligns with recent Pattern Recognition work on ranking, local kernels, graph selection, dynamic graphs, adaptive propagation, and molecular graph contrastive learning [56-61].

Step 11. Build pocket-wise and consensus leaderboards

For each pocket, emit a pocket-specific leaderboard,

$$B_{162}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{162}$$

Because ligand behavior may be pocket-dependent, the algorithm also constructs a consensus leaderboard. The consensus score is

$$B_{163}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{163}$$

where Agg is a predeclared aggregation rule.

A mean consensus is

$$B_{164}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{164}$$

A rank-product consensus is

$$B_{165}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{165}$$

A Borda-style consensus is

$$B_{166}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{166}$$

The aggregation rule must be logged in $\mathcal{M}$. This prevents post hoc selection of the most favorable pocket result.

The importance of pocket dependence is visible in SI Appendix I Figure 7, SI Appendix I Figure 8, and SI Appendix I Figure 9, where different docking engines, pocket choices, and component-energy patterns produce related but nonidentical rankings. Algorithm 1 treats this variability as information, not as noise to be hidden.

Step 12. Emit HSX dashboard outputs and provenance appendix when applicable

If any processed candidates match the HSX reference set, emit the HSX Mega-Core dashboard and attach Appendix O-HSX with the preferred IUPAC strings as a provenance-only table.

This step is documentary, not inferential. The HSX dashboard does not modify ranks unless its descriptors are explicitly included in the logged scoring protocol. Its purpose is to support traceability, inspection, and chemistry-safe comparison.

The HSX and HMC representations are interpreted as numerical reference architectures, not as atom-exact mega-scaffolds or synthesis prescriptions. This distinction is important. In the Hyper Mega Core topology map, nodes encode pharmacophore-role modules such as

$$B_{167}(P,L) = \mathrm{rDFT\ SDFT} + \mathrm{rQMC\ SQMC} + \mathrm{SMQW} - \Pi\_\mathrm{diag} \tag{167}$$

Edges encode descriptor-space compatibility, and thin teleport/refinement links encode kernel-space mapping for ranking and clustering. This interpretation is consistent with SI Appendix I Figure

3 and SI Appendix I Figure 4, where the HSX consensus references support audit-forward reporting rather than synthetic fusion claims.

The chemistry-safe abstraction layer is compatible with RDKit [18], SELFIES [22], grammar VAE [23], and junction-tree VAE [24]. It is used for representation and traceability, not for asserting synthesizability or therapeutic validity.

Step 13. Emit biology-facing persistence, MBHA, and QSAR-ready diagnostics when enabled

When biology-facing reporting is enabled, Algorithm 1 may compute recurrent interaction statistics. The normalized occupancy statistic is

$$\mathrm{B}_{168}(\mathrm{P,L}) = \mathrm{rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag} \tag{168}$$

Here c denotes an interaction class, S denotes receptor or structural states, T denotes time or sampling indices, and $I_{c,s,t}^{(k)}$ indicates whether interaction class c is observed for candidate k under state s and index t. This statistic identifies persistent interaction events rather than single-pose artifacts.

If MBHA-style search guidance is enabled, the update rule is

$$\mathrm{B}_{169}(\mathrm{P,L}) = \mathrm{rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag} \tag{169}$$

where x is the current candidate state, X_g is the elite or guiding attractor state, and $\odot$ denotes elementwise mixing. This update is interpreted as a controlled search heuristic, not as a chemical fusion operation.

A biology-facing composite score may be written as

$$\mathrm{B}_{170}(\mathrm{P,L}) = \mathrm{rDFT\ SDFT + rQMC\ SQMC + SMQW - \Pi\_diag} \tag{170}$$

The terms correspond to persistence support, docking-energy contribution, a penalty or regularizer, an interpretable spectral-SAR/QSAR hook, and an action-cost or stability penalty. All weights must be fixed or logged.

This biology-facing form is shown in SI Appendix I Figure 4. It extends the same audit-first logic to interaction persistence, stability filtering, and QSAR-ready interpretation. It does not convert the ranking output into claims of binding affinity, biological efficacy, or clinical performance.

Modern upstream systems such as AlphaFold [53], AlphaFold 3 [54], and DiffDock [55] may be used to provide structures, interaction hypotheses, or docking-oriented inputs. However, Algorithm 1 treats such outputs as upstream computational evidence, not as experimental validation.

Step 14. Finalize manifest and export audit outputs

Finalize the manifest

$$\mathcal{C}_{171}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{171}$$

and export all pocket-wise leaderboards, consensus leaderboards, branch diagnostics, rank-shift tables, topology flags, dashboard outputs, and protocol logs.

The exported package includes:

$$\mathcal{C}_{172}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{172}$$

and all branch flags.

The final export is therefore a structured audit archive, not merely a ranked table. The pocket–ligand comparability remains anchored by

$$\mathcal{C}_{173}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{173}$$

and

$$\mathcal{C}_{174}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{174}$$

Unsupported promotions trigger topology-disagreement flags through

$$\mathcal{C}_{175}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{175}$$

and unreliable branch outputs are penalized through

$$\mathcal{C}_{176}(\mathrm{P,L}) = \mathrm{f}_{\mathrm{rank}}(\mathrm{S}\theta, \bar{\mathrm{K}}, \Phi\mathrm{DFT}, \Phi\mathrm{QMC}, \mathrm{U(t)}, \Delta\mathrm{R}) \tag{176}$$

The final exported workflow is the one summarized in SI Appendix I Figure 1, SI Appendix I Figure 2, SI Appendix I Figure 5, and SI Appendix I Figure 6. Together, these figures show that

doi:10.20944/preprints202605.2096.v1

GLYBATOMAQ™ is a fused and audit-forward ranking system in which rank movement is attributable to stable, inspectable evidence rather than to any single branch alone.

Expanded pseudocode representation

The following pseudocode summarizes the algorithmic implementation.

Algorithm 1. GLYBATOMAQ™ quantum-enriched re-ranking and audit workflow

```
Input:
    Pocket set P
    Ligand sets L_P
    Base score S_theta(P,L)
    Docking outputs D(P,L)
    Operator descriptors Z_P, Z_L
    DFT protocol bundle Theta_DFT
    QMC protocol bundle Theta_QMC
    MQWalk protocol bundle Theta_MQW
    Fusion weights W
    Reliability rules R
    Shortlist size K

Output:
    Pocket-wise leaderboards B_P
    Consensus leaderboard B_cons
    Rank-shift diagnostics DeltaR
    Branch diagnostics
    MQWalk topology flags
    Optional HSX dashboard outputs
    Manifest M

0. Initialize manifest M
    Record receptor structures, pocket definitions, ligand sources,
    conformer settings, docking settings, descriptor settings,
    DFT/QMC/MQWalk settings, fusion weights, and reporting rules.

1. For each pocket P in pocket set:
        Compute or retrieve S_theta(P,L) for all L in L_P.
        Define L_top(P) = TopK(L_P; S_theta(P,.)).
        Assign R_0(P,L) on L_top(P).

2. For each ligand L in L_top(P):
        Assemble O(P,L):
            ligand identity
            optional SMILES/SELFIES
            conformer set
            selected geometry chi_star(P,L)
            docking pose provenance
            docking output D(P,L)
            base score S_theta(P,L)

3. Construct PSD operators:
        rho_P = (1/n_P) Z_P^T Z_P
        rho_L = (1/n_L) Z_L^T Z_L
```

```
          K(P,L) = Tr(rho_P rho_L)
          K_bar(P,L) = Tr(rho_P rho_L) /
                             (sqrt(Tr(rho_P^2)Tr(rho_L^2)) + epsilon)

4. Apply shared conditioning:
          For each feature f_k:
                z_k = (f_k - median(f_k)) / (MAD(f_k) + epsilon)
                s(alpha) = exp(-alpha log N_A)
                f_hat_k = tanh(z_k s(alpha))

5. DFT branch:
          If enabled:
                Run DFT on chi_star(P,L).
                Extract Phi_DFT(P,L).
                Compute S_DFT(P,L) = beta_DFT^T Phi_DFT(P,L).
                Compute r_DFT(P,L).
                Record DFT flags.
          Else:
                S_DFT = 0
                r_DFT = 0
                flag = DFT_NOT_RUN

6. QMC branch:
          If enabled:
                Estimate E_QMC_hat, sigma_QMC, CIwidth, ESS.
                Compute S_QMC = -E_QMC_hat
                                      - lambda_sigma sigma_QMC
                                      - lambda_c CIwidth.
                Compute r_QMC(P,L).
                Record QMC flags.
          Else:
                S_QMC = 0
                r_QMC = 0
                flag = QMC_NOT_RUN

7. MQWalk branch:
          Construct W = [[0, K_bar], [K_bar^T, 0]].
          A = D^(-1/2) W D^(-1/2).
          A_gamma = (1-gamma)A + gamma(1/n)11^T.
          L_tilde = s(alpha)(I - A_gamma).
          U(t) = exp(-i t L_tilde).
          p_j(t) = |<j|U(t)|psi_0>|^2.
          Compute MQWscore(P,L).
          Compute Delta_topo(P,L).
          Compute S_MQW(P,L).

8. Diagnostic penalty:
          Compute Pi_diag(P,L) from:
                DFT flags
                QMC uncertainty
                topology disagreement
```

```
            convergence failures
            missing provenance

9. Fusion:
        S_final(P,L) =
            w_0 S_theta(P,L)
            + w_DFT r_DFT S_DFT(P,L)
            + w_QMC r_QMC S_QMC(P,L)
            + w_MQW S_MQW(P,L)
            - Pi_diag(P,L)

10. Post-refinement audit:
        Assign R_+(P,L) by sorting S_final(P,L).
        DeltaR(P,L) = R_0(P,L) - R_+(P,L).
        Compute Overlap@k(P).
        Compute Spearman correlation between R_0 and R_+.
        Emit pocket-wise leaderboard B_P.

11. Consensus:
        Compute S_cons(L) = Agg_P S_final(P,L).
        Emit consensus leaderboard B_cons.

12. HSX and dashboard outputs:
        If HSX candidates are present:
            Emit HSX dashboard.
            Attach provenance-only Appendix O-HSX.

13. Biology-facing diagnostics:
        If enabled:
            Compute persistence P_c^(k).
            Compute MBHA-guided search summaries.
            Compute QSAR-ready composite diagnostics.

14. Finalize:
        Export leaderboards, branch diagnostics, topology flags,
        rank shifts, manifest M, and all replay logs.
```

Mathematical role of each algorithmic block

The algorithm can be understood as the following sequence of controlled transformations:

$$\mathcal{C}_{177}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{177}$$

Here,

$$\mathcal{C}_{178}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{178}$$

is the operator-kernel graph,

$$\mathcal{C}_{179}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{179}$$

is the electronic and uncertainty evidence bundle, and

$$\mathcal{C}_{180}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{180}$$

is the topology-validation bundle.

The rank movement

$$\mathcal{C}_{181}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{181}$$

is the final object to be interpreted.

doi:10.20944/preprints202605.2096.v1

This formulation is important because it makes the mathematical-chemistry contribution explicit. The algorithm does not merely generate a new score. It explains how a candidate moved, why it moved, and whether that movement is reliable.

Functional role of the layered quantum-geometric architecture

The algorithm contains several layers, but each layer has a defined failure mode that it addresses. The docking layer addresses structural plausibility under a fixed computational protocol [6-16]. The PSD operator-kernel layer addresses instability caused by comparing heterogeneous ligand and pocket descriptors without a controlled geometry [33,36]. The DFT branch addresses cases where chemically meaningful electronic differences are not captured by docking alone [44-47]. The QMC branch addresses uncertainty in energetic refinement and prevents overconfident interpretation of noisy estimates [48,49]. The MQWalk branch addresses topology-discordant promotions, where a candidate is promoted by one branch but unsupported by the similarity neighborhood [50-52]. The diagnostic penalty layer addresses convergence failure, inflated uncertainty, topology disagreement, and missing provenance. The rank-audit layer addresses the final interpretability problem: whether the leaderboard changed in a way that can be explained. Thus, the complexity is not introduced to make the workflow appear advanced. It is introduced to prevent five common errors in computational screening:

$$\mathcal{C}_{182}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{182}$$

Algorithm 1 is designed specifically to avoid these errors.

Conservative interpretation of the final output

The final output of Algorithm 1 is a reproducible prioritization hypothesis. It is not a claim of experimental binding, therapeutic efficacy, selectivity, toxicity, or clinical value.

A candidate with strong output support satisfies the following evidence chain:

$$\mathcal{C}_{183}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{183}$$

A candidate with weak or questionable support may still appear in the leaderboard, but it will be accompanied by diagnostic flags such as

$$\mathcal{C}_{184}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{184}$$

This output style is intentionally audit-forward. It allows the reader to distinguish a stable promotion from an unstable or branch-specific artifact.

Final expanded algorithm statement

Algorithm 1 implements GLYBATOMAQ™ as a replayable mathematical-chemistry ranking transformation. It begins with fixed-protocol docking and pattern-recognition outputs, constructs pocket-wise shortlists, embeds pocket and ligand information into positive-semidefinite operator summaries, evaluates deterministic DFT descriptors and uncertainty-aware QMC energies, validates rank promotions through MQWalk propagation on the normalized operator-overlap graph, and fuses all evidence through a reliability-gated score. The final product is a pocket-wise and consensus leaderboard accompanied by rank shifts, branch contributions, topology flags, and a complete manifest.

In compact form, the algorithm computes

$$\mathcal{C}_{185}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{185}$$

and audits the resulting movement through

$$\mathcal{C}_{186}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{186}$$

with topology controlled by

$$\mathcal{C}_{187}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{187}$$

and MQWalk validation governed by

$$\mathcal{C}_{188}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{188}$$

where

$$\mathcal{C}_{189}(P,L) = f_{\mathrm{rank}}(S\theta, \bar{K}, \Phi DFT, \Phi QMC, U(t), \Delta R) \tag{189}$$

These equations make the algorithm's contribution explicit: GLYBATOMAQ™ is not a non-functional combination of quantum-sounding modules. It is a structured ranking audit in which electronic refinement, stochastic uncertainty, and topology consistency are integrated to explain and constrain ligand-prioritization decisions.

2.1.10. Quantum-Geometric 2D Proof of MQWalk-MBHA-Guided Pharmacophore Assembly

2.1.10.1. Scope of the Quantum-Geometric Construction

The preceding sections define GLYBATOMAQ™ as a rank-centric, chemistry-safe re-ranking framework. The present subsection strengthens that formulation by expressing the two-dimensional HMC assembly process as a quantum-geometric optimization problem on a constrained pharmacophore manifold. The construction is mathematical and algorithmic. It does not assert coherent quantum transport in GLIPR1, atom-exact fusion of all modules, a synthetic route, binding affinity, biological efficacy, or therapeutic activity. Instead, it provides a formal certificate showing that GLYBATOMAQ-HSX candidates are generated as ordered pharmacophore paths whose planar geometry, operator compatibility, graph topology, curvature regularity, and uncertainty-aware rank evidence are jointly audited [23,24,30,31,33,36,44-52].

2.1.10.2. HMC Role Grammar as a Constrained 2D Pharmacophore Manifold

Let the Hyper Mega Core grammar be represented by six ordered pharmacophore-role classes. The first five roles define the mandatory assembly backbone, whereas the sixth role represents an optional terminal cap or tail used only when the candidate design requires it:

$$\mathrm{GHMC} = (M_1, M_2, M_3, M_4, M_5, M_6) \qquad (190)$$

The roles correspond to recognition nucleus, linker, interaction node, spacer, polarity/solubility module, and optional tail/cap. An admissible GLYBATOMAQ-HSX candidate is therefore an ordered typed chain rather than an arbitrary mixture of fragments:

$$L = (m_1, m_2, m_3, m_4, m_5, m_6),\ m_i \in M_i \qquad (191)$$

Each module is embedded in a planar pharmacophore chart by an anchor coordinate, an orientation, and a compact footprint. Thus the local 2D state of module i is

$$q_i = (p_i, \varphi_i, G_i),\ p_i = (x_i,y_i)\in\mathbb{R}^2,\ \varphi_i\in S^1 \qquad (192)$$

The complete planar assembly is the ordered object

$$L_2D = q_{ii} = 1^6,\ e_i = [p_i,p_i + {}_1],\ i = 1,\ldots,5 \qquad (193)$$

The feasible pharmacophore manifold is defined by role order, edge-length bounds, angular bounds, orientation compatibility, and non-overlap of non-adjacent footprints:

$$\boldsymbol{\mathcal{F}}_2D = L_2D\text{: } d_i,\min \leq \| p_i + {}_1 - {}_{pi} \| \leq d_i,\max,\ \theta_i,\min \leq \theta_i \leq \theta_i,\max,\ |\varphi_i + {}_1 - \varphi_i| \leq \eta_i \qquad (194)$$

This manifold excludes disconnected, collapsed, and topologically incoherent arrangements while preserving the chemistry-safe interpretation of the HMC. The role grammar specifies which pharmacophore functions may be adjacent; the 2D feasibility constraints specify how those functions may be geometrically realized.

2.1.10.3. Operator-Valued Anchors and Normalized Quantum-Geometric Overlap

To couple the planar pharmacophore chain to the operator-kernel layer of the manuscript, each module is assigned a local feature matrix Z_i and a positive-semidefinite operator. This construction parallels the pocket-ligand PSD representation used throughout the GLYBATOMAQ™ ranking layer:

$$\varrho_i = Z_i^T Z_i / \mathrm{Tr}(Z_i^T Z_i),\ \varrho_i\succeq 0,\ \mathrm{Tr}(\varrho_i) = 1 \qquad (195)$$

The module operator is therefore a normalized density-like descriptor object. It is not interpreted as a physical molecular density matrix; it is an auditable quantum-geometric representation of local pharmacophore information. Compatibility between adjacent modules is measured by normalized trace overlap,

$$\bar{K}_{ij} = \mathrm{Tr}(\varrho_i\ \varrho_j)/(\surd[\mathrm{Tr}(\varrho_i^2)\ \mathrm{Tr}(\varrho_j^2))+\varepsilon) \qquad (196)$$

A role-adjacent edge is admitted only when descriptor-space overlap and geometric feasibility are satisfied simultaneously:

$$(m_i,m_i + {}_1)\in E \Leftrightarrow [\bar{K}_i,i+1 \geq \tau_i] \wedge [L_2D\in\boldsymbol{\mathcal{F}}_2D] \tag{197}$$

The connection step is therefore a kernel-permission rule rather than a freehand scaffold-merge operation. It couples local 2D pharmacophore geometry to a PSD operator overlap, so that each edge has both geometric and descriptor-space justification [33,36].

2.1.10.4. Quantum-Geometric Metric, Distance, and Curvature on the Assembly Chart

The quantum-geometric enrichment is obtained by treating the parameterized assembly state x as a point on a constrained descriptor manifold. Let x collect module anchors, orientations, edge lengths, angles, and branch descriptors:

$$x = (p_1,\ldots,p_6,\ \varphi_1,\ldots,\varphi_6,\ l_1,\ldots,l_5,\ \theta_2,\ldots,\theta_5,\ \zeta),\ l_i = \Vert p_i + {}_1 - {}_{pi}\Vert \tag{198}$$

A small perturbation of x induces a perturbation of the operator $\varrho(x)$. The local quantum-geometric metric is defined by the real part of the Hilbert-Schmidt pullback metric,

$$g_{a\beta}({}_x) = \mathrm{Re}\ \mathrm{Tr}[(\partial_a\varrho(x))(\partial_\beta\varrho(x))] \tag{199}$$

This metric quantifies how strongly a small planar or descriptor-space deformation changes the operator representation. For pure-state embeddings $|\psi(x)>$, the corresponding Fubini-Study form is

$$ds^2FS = \langle d\psi|d\psi\rangle - |\langle\psi|d\psi\rangle|^2 \tag{200}$$

For mixed PSD descriptors, a Bures-style distance supplies a conservative dissimilarity measure between two pharmacophore operators:

$$D\_B^2(\varrho_i,\varrho_j) = 2[1 - \mathrm{Tr}((\varrho_i{}^1/2\varrho_j\varrho_i{}^1/2)^1/2)] \tag{201}$$

The connection rule can therefore be sharpened by combining overlap with quantum-geometric distance:

$$(m_i,m_i + {}_1)\in E \Leftrightarrow \bar{K}_i,i+1 \geq \tau_i \text{ and } D\_B(\varrho_i,\varrho_i + {}_1) \leq \delta_i \tag{202}$$

Curvature is introduced to penalize abrupt descriptor-space bending that is not supported by the pharmacophore chain. If g_ab defines the local metric, then the induced geodesic energy of a candidate path γ is

$$E_{9eo}(\gamma) = \int_0^1 \dot{\gamma}^{a(t)g}{}_{a\beta}(\gamma(t))\dot{\gamma}^{b(t)dt} \tag{203}$$

A candidate with low geometric energy changes smoothly across module roles, whereas a high-energy path indicates an abrupt operator-space transition. This provides a quantum-geometric regularizer for the HMC assembly.

2.1.10.5. Phase/Gauge Consistency and Berry-Type Assembly Curvature

Because operator embeddings may be represented by local amplitudes or normalized descriptor states, the absolute phase of a local representation is not decision-relevant. What matters is whether the relative change between neighboring modules is consistent along the assembled path. A local gauge connection can be written as

$$A_a({}_x) = \mathrm{Im}\langle\psi(x)|\partial_a\psi(x)\rangle \tag{204}$$

The corresponding curvature two-form is

$$\Omega_{a\beta}({}_x) = \partial_a A_\beta({}_x) - \partial_\beta A_a({}_x) \tag{205}$$

For a closed comparison loop C in the HMC graph, the accumulated geometric phase surrogate is

$$\Gamma(C) = \oint C\ A_a\ dx^a \tag{206}$$

In GLYBATOMAQ™, this term is not used as a physical Berry phase claim. It is used as a graph-geometric consistency diagnostic: large loop inconsistency indicates that the proposed assembly changes representation abruptly or depends strongly on path choice. The corresponding penalty is

$$\Pi_{curv}(L) = \lambda\Omega\int\gamma\Vert\Omega_{a\beta}(\gamma(t))\Vert F^2dt + \lambda\Gamma|\Gamma(CL)| \tag{207}$$

This term makes the proof more stringent than a simple distance-and-angle construction: an accepted candidate should be geometrically feasible, operator-compatible, and locally smooth under the induced quantum-geometric connection.

2.1.10.6. Layered HMC Graph, Quantum Walk Support, and Topology Certificate

Define the layered HMC graph $H=(V,E)$, with $V=V_1\cup\ldots\cup V_6$ and V_i containing admissible modules of role M_i. A generated ligand is a six-layer path

$$v_1 \in V_1, \ldots, v_6 \in V_6, (v_i, v_i + {}_1) \in E \text{ for } i = 1, \ldots, 5 \tag{208}$$

The planar and operator-geometric admissibility certificate is therefore

$$L \in \boldsymbol{\mathcal{F}}_2 D \cap \mathrm{Path}(H) \cap \mathrm{QG}(H) \tag{209}$$

where QG(H) denotes the set of paths satisfying PSD-overlap, Bures-distance, and curvature-consistency constraints. MQWalk supplies the topological support audit. With degree-normalized adjacency A, teleport-stabilized adjacency A_γ, and Avogadro-conditioned Laplacian $\tilde{L}$,

$$U(t) = \exp(-it\tilde{L}), \; p_j({}_t) = |\langle j | U(t) | \psi_0 \rangle|^2 \tag{210}$$

The path-support functional is defined as the time-averaged propagated mass over the candidate path:

$$S_{\mathrm{path}}(L) = T^{-1} \int_0^T \Sigma_j \in L p_j({}_t)^d{}_t \tag{211}$$

A topology-supported candidate must retain sufficient propagated mass in its compatible HMC neighborhood:

$$S_{\mathrm{path}}(L) \geq \kappa_{\mathrm{path}} \tag{212}$$

Thus, MQWalk does not create module connections. It tests whether the connections admitted by the quantum-geometric HMC graph remain coherent under propagation on the operator-overlap topology [50-52].

2.1.10.7. MBHA as Quantum-Geometric Projection Dynamics

MBHA generates new candidate geometries by mixing the current state with an elite attractor while preserving the modular state representation:

$$x_{\mathrm{ra}} w = C_1 \odot X_9 + C_2 \odot x \tag{213}$$

The enriched projection step maps the raw update back into the feasible quantum-geometric domain:

$$x_{\mathrm{ne}} w = \Pi \boldsymbol{\mathcal{F}}_2 D \cap \mathrm{Path}(H) \cap \mathrm{QG}(H)(x_{\mathrm{ra}} w) \tag{214}$$

Equivalently, the projection can be defined as the minimizer of a composite geometric discrepancy:

$$\Pi\_\mathrm{QG}({}_\gamma) = \arg\min{}_x \in \boldsymbol{\mathcal{F}}_2 D \cap \mathrm{Path}(H) \cap \mathrm{QG}(H) [\|x-y\|^2 + \lambda B \Sigma_i D\_B^2(\varrho_i, \varrho_i + {}_1) + \lambda_9 E_{9_{\mathrm{eo}}}(\gamma_x) + \Pi_{\mathrm{curv}}({}_x)] \tag{215}$$

This equation clarifies the role of MBHA in the enriched proof. MBHA is not an unrestricted generator of fused molecular structures; it is a projected search operator on the feasible HMC pharmacophore manifold. The projection retains role order, 2D feasibility, PSD compatibility, quantum-geometric smoothness, and MQWalk topology support.

2.1.10.8. Reliability-Weighted Quantum-Geometric Objective and Theorem

The DFT and QMC branches refine the generated path by attaching deterministic electronic descriptors and uncertainty-aware energetic evidence to an already admissible HMC assembly [44-49]. The enriched final objective is

$$S\_\mathrm{QG}^+ (P,L) = w_0 S\theta + w D_r DSD + w Q_r QSQ + wMSM + w_{\mathrm{path}} S_{\mathrm{path}} - \Pi\_\mathrm{diag} - \Pi_{\mathrm{curv}} - \Pi\_\mathrm{QG} \tag{216}$$

Here Π_QG collects distance, metric, and projection-residual penalties not already represented in Π_diag or Π_curv. The final GLYBATOMAQ-HSX candidate is obtained by

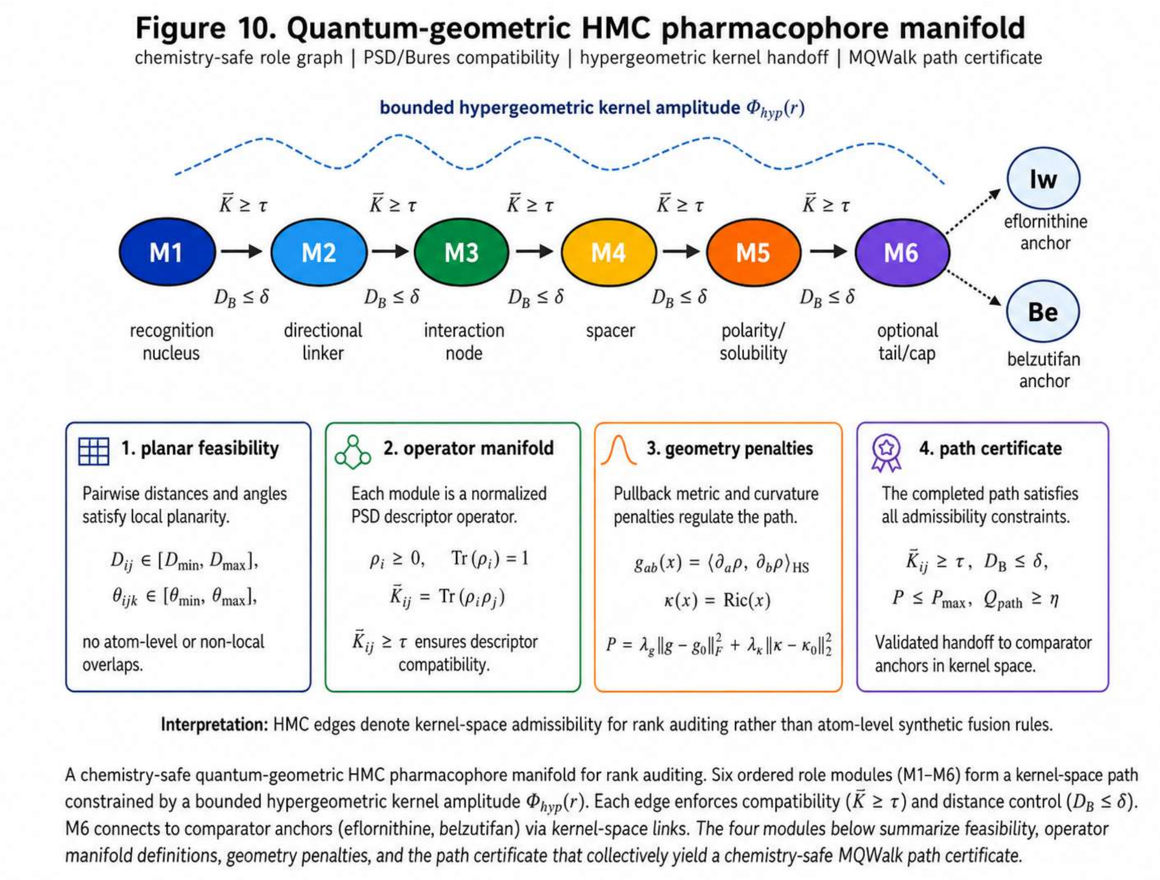
$$L^* = \arg\max_{L\in\mathcal{F}_2D\cap \mathrm{Path}(H)\cap \mathrm{QG}(H)} S\_QG^{+} (P,L) \tag{217}$$

Theorem. Under the HMC grammar, PSD operator mapping, quantum-geometric metric, Bures-distance edge rule, curvature-consistency penalty, MQWalk topology support, MBHA projection, and reliability-weighted DFT/QMC fusion, GLYBATOMAQ-HSX ligand generation is a constrained quantum-geometric graph optimization problem rather than an arbitrary scaffold-fusion procedure.

Proof. The HMC grammar fixes the pharmacophore roles and their order. The planar feasibility set fixes the allowable 2D anchor, angle, orientation, and non-overlap geometry. PSD operator mapping assigns each module a normalized descriptor operator, and trace overlap plus Bures-style distance determines which role-adjacent pairs may be connected. The metric g_ab, geodesic energy, and curvature penalty constrain abrupt descriptor-space deformation and path-dependent inconsistency. MQWalk then tests whether the assembled path is supported by the operator-overlap graph. MBHA proposes new candidates only through projection into the feasible quantum-geometric domain. Finally, DFT, QMC, MQWalk, diagnostic, curvature, and projection terms are combined in the reliability-weighted score. Therefore any accepted candidate has an explicit assembly certificate:

Certificate(L) = role order, $\mathcal{F}_2$D, PSD overlap, D_B bound, curvature consistency, MQWalk support, MBHA projection, DFT/QMC audit (218)

This certificate proves that GLYBATOMAQ-HSX outputs are reproducible 2D quantum-geometric pharmacophore hypotheses for GLIPR1-oriented ranking. They remain computational design hypotheses and should not be interpreted as experimentally validated ligands, therapeutic structures, or synthetic prescriptions.



Interpretation: HMC edges denote kernel-space admissibility for rank auditing rather than atom-level synthetic fusion rules.

A chemistry-safe quantum-geometric HMC pharmacophore manifold for rank auditing. Six ordered role modules (M1–M6) form a kernel-space path constrained by a bounded hypergeometric kernel amplitude $\Phi_{hyp}(r)$. Each edge enforces compatibility ($\bar{K} \geq \tau$) and distance control ($D_B \leq \delta$). M6 connects to comparator anchors (eflornithine, belzutifan) via kernel-space links. *The four modules below summarize feasibility, operator manifold definitions, geometry penalties, and the path certificate that collectively yield a chemistry-safe MQWalk path certificate.*

Figure 10. Quantum-geometric HMC pharmacophore manifold and chemistry-safe assembly certificate. The HMC is shown as an ordered pharmacophore-role manifold rather than an atom-exact fused scaffold. Modules are admitted only when planar feasibility, normalized PSD operator overlap, bounded hypergeometric kernel compatibility, Bures/Fubini-Study distance control, curvature smoothness, MQWalk path support, and

DFT/QMC reliability are simultaneously compatible. Comparator anchors are linked only in kernel space, preserving chemistry-safe interpretation.

2.1.11. Quantum-Geometry Reporting Module for the Original Research Workflow

This study includes a dedicated reporting module that separates quantum-geometric definitions from numerical claims. The module makes the mathematical contribution inspectable by linking each metric, distance, curvature, sensitivity, and topology term to a specific decision in the GLYBATOMAQ™ rank-audit workflow. The figures and tables are cited in the main text as part of the integrated method and results narrative.

The central object is a normalized PSD descriptor operator, rho(x), attached to a ligand, pocket, residue-contact neighborhood, or HMC role module. Local perturbations in x define a pullback metric g_ab(x), pairwise comparisons define overlap and Bures-style distances, path continuity is controlled by geodesic energy and curvature penalties, and MQWalk evaluates whether the resulting candidate path retains support in the operator-overlap graph. These quantities are interpreted as computational geometry diagnostics, not as direct physical observables.

Figures 10-14 and Tables 12-15 provide the visual and tabular core of the quantum-geometric reporting architecture. Figure 10 defines the chemistry-safe HMC manifold and assembly certificate; Figure 11 links bounded hypergeometric operator amplitudes to the DFT/QMC/MQWalk audit circuit and SI Appendix I biological anchors; Figure 12 provides the proxy-overlay generalization surface for chemical-energy, MMP3, and candidate-level support; Figure 13 defines the normalized diagnostic reporting matrix; and Figure 14 visualizes MQWalk topology validation as propagated mass over the operator-overlap graph. Tables 12-15 give the corresponding object definitions, candidate reporting fields, cross-reference map, and reproducibility checklist.

Figure 11. Quantum-hypergeometric DFT/QMC/MQWalk audit circuit

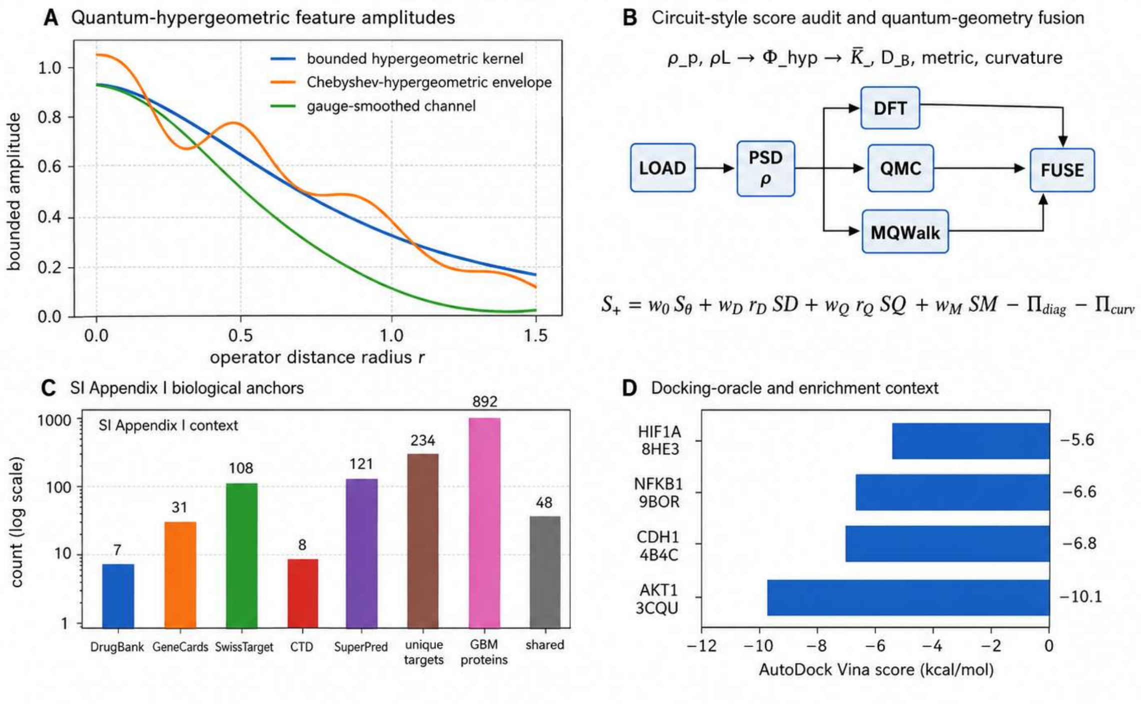


Figure 11. Bounded hypergeometric and Chebyshev-hypergeometric feature channels (A) are fused with DFT/QMC/MQWalk branches in a circuit-style audit pipeline (B) to produce a composite quantum-geometry score. Biological anchors from SI Appendix I (C) and docking-oracle context (D) provide target-harvesting and enrichment support for contextual interpretation.

doi:10.20944/preprints202605.2096.v1

Figure 11. Quantum-hypergeometric DFT/QMC/MQWalk audit circuit. Bounded hypergeometric and Chebyshev-hypergeometric feature channels define operator amplitudes that are passed through PSD encoding, DFT descriptor scoring, QMC uncertainty-aware refinement, MQWalk topology validation, and reliability-weighted fusion. Biological anchors from SI Appendix I Figure 2 are included as target-harvesting and docking-oracle context, including 234 vortioxetine targets, 892 glioblastoma proteins, 48 shared targets, and AKT1/CDH1/NFKB1/HIF1A docking-score patterns.

Figure 12. Proxy-overlay generalization of quantum-geometric screening evidence

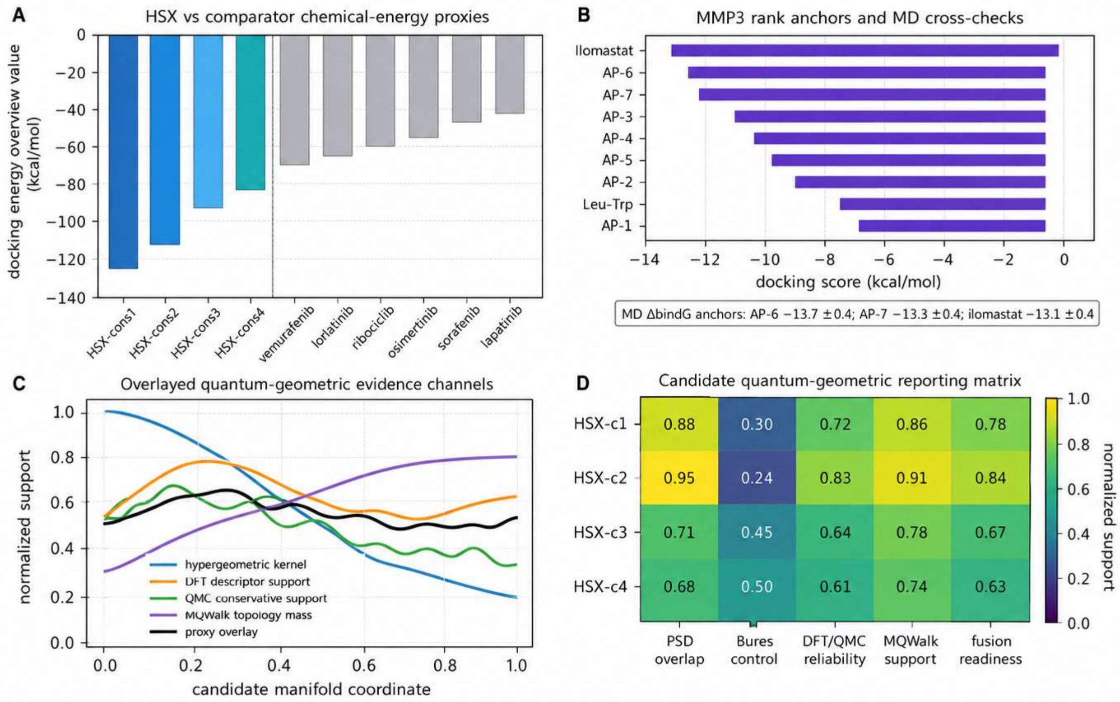


Figure 12. Proxy-overlay generalization of the quantum-geometric GLYBATOMAQ audit. HSX chemical-energy proxies (A), MMP3 inhibitor ranking anchors with MD cross-checks (B), and candidate-level quantum-geometric support channels (C) are projected onto a common operator-geometric rank-audit surface. The candidate quantum-geometric reporting matrix (D) combines PSD overlap, Bures control, DFT/QMC reliability, MQWalk support, and fusion readiness using biological and chemical context from SI Appendix I.

Figure 12. Proxy-overlay generalization of the quantum-geometric GLYBATOMAQ audit. HSX chemical-energy proxies, MMP3 inhibitor ranking anchors, DFT/QMC/MQWalk support traces, and a candidate-level quantum-geometric reporting matrix are projected onto a common operator-geometric surface. The biological and chemical anchors summarize SI Appendix I Figs. 1 and 5 and are used as contextual support for rank-audit interpretation, not as experimental efficacy claims.

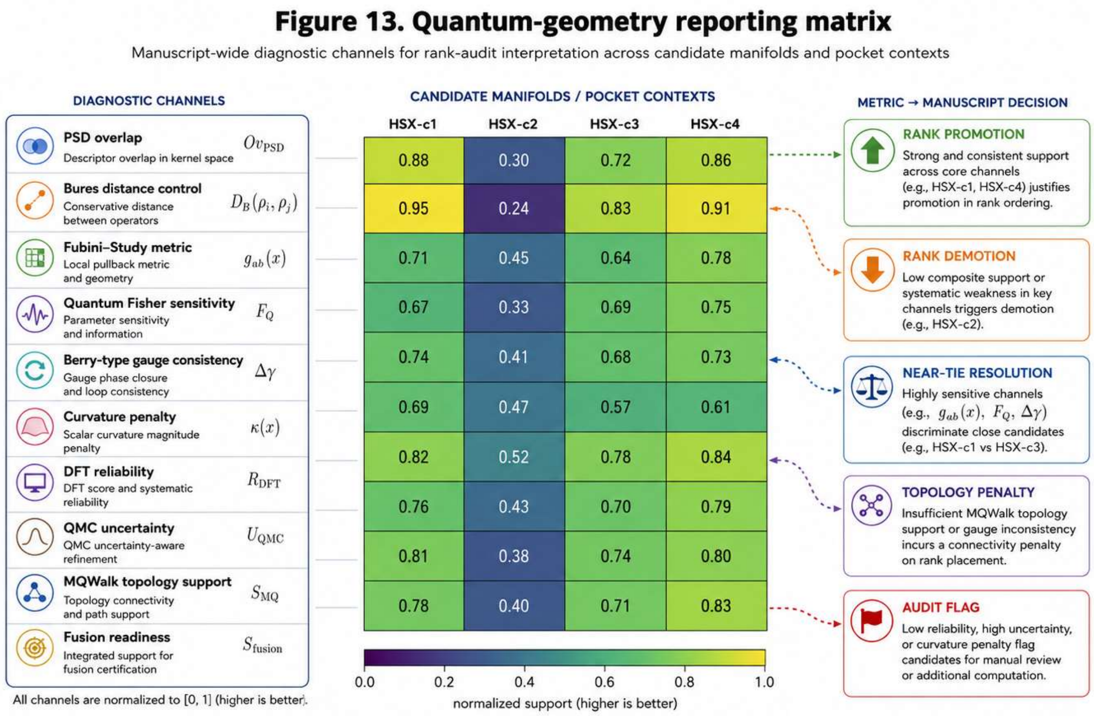


Figure 13. Quantum-geometry reporting matrix

Manuscript-wide diagnostic channels for rank-audit interpretation across candidate manifolds and pocket contexts

DIAGNOSTIC CHANNELS			HSX-c1	HSX-c2	HSX-c3	HSX-c4
PSD overlap	Descriptor overlap in kernel space	Ov_{PSD}	0.88	0.30	0.72	0.86
Bures distance control	Conservative distance between operators	$D_B(\rho_i, \rho_j)$	0.95	0.24	0.83	0.91
Fubini–Study metric	Local pullback metric and geometry	$g_{ab}(x)$	0.71	0.45	0.64	0.78
Quantum Fisher sensitivity	Parameter sensitivity and information	F_Q	0.67	0.33	0.69	0.75
Berry-type gauge consistency	Gauge phase closure and loop consistency	$\Delta\gamma$	0.74	0.41	0.68	0.73
Curvature penalty	Scalar curvature magnitude penalty	$\kappa(x)$	0.69	0.47	0.57	0.61
DFT reliability	DFT score and systematic reliability	R_{DFT}	0.82	0.52	0.78	0.84
QMC uncertainty	QMC uncertainty-aware refinement	U_{QMC}	0.76	0.43	0.70	0.79
MQWalk topology support	Topology connectivity and path support	S_{MQ}	0.81	0.38	0.74	0.80
Fusion readiness	Integrated support for fusion certification	S_{fusion}	0.78	0.40	0.71	0.83

RANK PROMOTION
Strong and consistent support across core channels (e.g., HSX-c1, HSX-c4) justifies promotion in rank ordering.

RANK DEMOTION
Low composite support or systematic weakness in key channels triggers demotion (e.g., HSX-c2).

NEAR-TIE RESOLUTION
Highly sensitive channels (e.g., $g_{ab}(x)$, F_Q, $\Delta\gamma$) discriminate close candidates (e.g., HSX-c1 vs HSX-c3).

TOPOLOGY PENALTY
Insufficient MQWalk topology support or gauge inconsistency incurs a connectivity penalty on rank placement.

AUDIT FLAG
Low reliability, high uncertainty, or curvature penalty flag candidates for manual review or additional computation.

Figure 13. Quantum-geometry reporting matrix. The matrix organizes the primary diagnostic channels used throughout the manuscript into a unified reporting schema. Values are normalized to [0, 1] (higher is better) for each candidate manifold or pocket context (HSX-c1 to HSX-c4). Channel definitions and exemplar expressions are given at left; decision mappings at right link metric behavior to manuscript actions including promotion, demotion, near-tie resolution, topology penalties, and audit flags. This figure provides the canonical reporting schema for cross-figure interpretation and should be cited in the main body where diagnostic channels are introduced.

Figure 13. Quantum-geometry reporting matrix. The matrix organizes PSD overlap, Bures distance control, Fubini-Study pullback geometry, quantum Fisher sensitivity, Berry-type gauge consistency, curvature penalty, DFT reliability, QMC uncertainty, MQWalk topology support, and fusion readiness into a manuscript-wide diagnostic reporting schema. The decision boxes link normalized metrics to rank promotion, demotion, near-tie resolution, topology penalties, and audit flags.

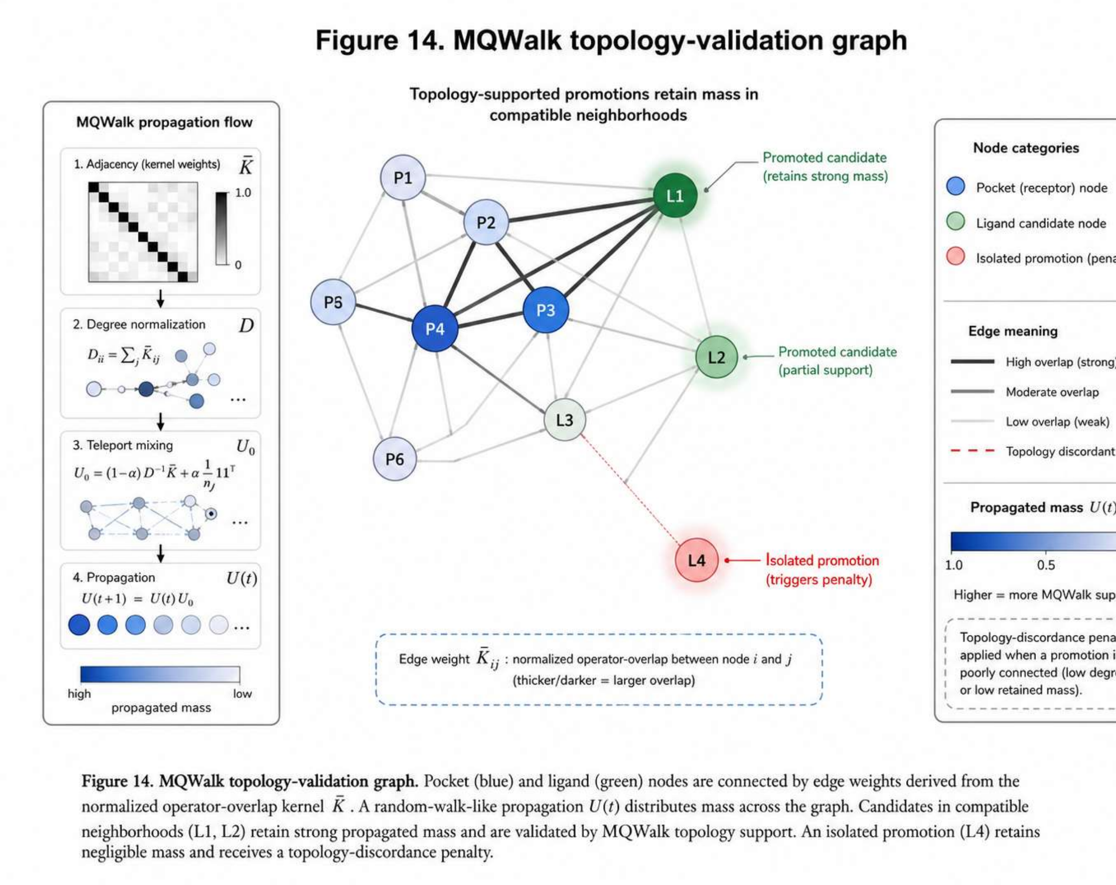


Figure 14. MQWalk topology-validation graph. Pocket and ligand nodes are connected by normalized operator-overlap weights. Topology-supported promotions retain propagated mass in compatible neighborhoods, whereas isolated promotions receive topology-discordance penalties. This visualization operationalizes the graph-consistency interpretation of MQWalk and should be read together with SI Appendix I Figs. 1B-C, 3, and 6.

Table 12. Quantum-geometric objects used in the GLYBATOMAQ™ article and their computational interpretation.

Quantum-geometric object	Formal role in GLYBATOMAQ™	Rank-audit interpretation	Primary citation in manuscript
Normalized PSD operator ρ	Trace-normalized descriptor state for pocket, ligand, residue, or HMC module	Places heterogeneous descriptors on a comparable operator manifold	Section 2.1.10; Figure 11
Bures-style distance D^B	Conservative distance between PSD descriptor operators	Prevents candidate paths from connecting incompatible modules solely by scalar score	Sections 2.1.10.4 and 2.1.11; Table 12
Fubini-Study / pullback metric $g_{a\beta}$	Local geometry induced by descriptor perturbations	Quantifies smoothness and deformation sensitivity of the pharmacophore path	Figure 12; Table 13
Quantum Fisher information-style sensitivity	Metric-related sensitivity of the descriptor state to local perturbations	Flags unstable descriptors or overly sharp rank changes	Figure 13

Berry-type curvature / gauge consistency	Path-dependence and phase/gauge inconsistency diagnostic	Penalizes abrupt or inconsistent descriptor-space bending	Figures 12 and 13
MQWalk propagated mass	Topology support under operator-overlap graph propagation	Tests whether DFT/QMC-supported promotions are neighboϱod-compatible	Figure 14

Table 13. Candidate-level quantum-geometry reporting matrix for GLYBATOMAQ-HSX leaderboards.

Reporting field	Symbol or output	Expected value type	Use in decision rule
PSD overlap	K̄(P,L)	0-1 or normalized scalar	Initial geometry-compatible similarity support
Bures distance	D_B(ϱ_i,ϱ_j)	non-negative scalar	Edge admissibility and path penalty
Geodesic energy	E_geo(gamma)	non-negative scalar	Smoothness penalty along HMC role path
Curvature penalty	Pi_curv	non-negative scalar	Gauge/curvature consistency penalty
QFI-style sensitivity	I_QG or local metric trace	normalized scalar/vector	Robustness of descriptor response
MQWalk support	S_path or S_MQW	normalized scalar	Topology-supported promotion evidence
Projection residual	Pi_QG	non-negative scalar	Distance from feasible quantum-geometric domain
Final audit class	supported / weak / failed	categorical flag	Publication-ready interpretation of rank movement

Table 14. Cross-reference map for the quantum-geometry figures and tables included in the article.

Figure/Table	Where cited in the manuscript	Purpose of citation
Figure 11	Introduction; Section 2.1.11; Discussion 4.2/4.7	Provides the overview map for the enriched quantum-geometry audit architecture
Figure 12	Section 2.1.10.4; Section 2.1.11; Discussion 4.3	Explains HMC/HSX feasibility as a constrained pharmacophore manifold
Figure 13	Results 3.11; Discussion 4.14	Standardizes matrix-style reporting of quantum-geometry descriptors
Figure 14	Methods 2.1.6; Results 3.8/3.11; Discussion 4.7	Shows how MQWalk topology validation constrains promoted candidates
Tables 12-13	Methods 2.1.11; Results 3.11	Define and operationalize the quantum-geometric descriptor panel
Tables 14-15	Discussion 4.14; declarations/reproducibility section	Ensure that every visual/table element is cited and auditable

Table 15. Reproducibility checklist for quantum-geometric GLYBATOMAQ™ reporting.

Checklist item	Minimum reporting requirement	Status in the manuscript
Protocol boundary	State docking is a ranking oracle, not affinity measurement	Included in Abstract, Methods, Discussion, and Limitations
Descriptor provenance	List feature source, normalization, and PSD construction	Included in Sections 2.1.2, 2.1.3, and Table 12

Quantum-geometry metrics	Report overlap, distance, metric, curvature, and projection terms	Defined in Section 2.1.10 and Tables 12-13
Uncertainty treatment	Report QMC variance/CI/ESS and reliability gate	Included in Sections 2.1.5, 2.1.9.5, and Results
Topology support	Report MQWalk mass, discordance, and graph support	Included in Sections 2.1.6, 2.1.10.6, Figure 14
Rank audit	Report pre-rank, post-rank, rank shift, penalties, and audit class	Included in Sections 2.1.7, 3.11, and Table 13
Interpretation limit	Separate computational hypotheses from biological or clinical claims	Included in Scope, Limitations, and Conclusions

2.1.12. Expanded Classical Mathematical Formalism for the Quantum-Geometric Rank-Audit Layer

The following equations collect the quantum-geometric core of GLYBATOMAQ™ in a classical display style suitable for journal reading. They are written as operational definitions rather than physical claims: the operators are normalized descriptor states, the distances and metrics are information-geometric controls, and the walk terms are topology diagnostics on the operator-overlap graph. The notation links directly to Figures 10-14 and Tables 12-15, which provide the graphical and tabular reporting schema used in the Results, Discussion, and Conclusion sections.

$$\varrho(x) = Z(x)\, Z(x)^T / \mathrm{Tr}[Z(x)\, Z(x)^T],\ \varrho(x) \succeq 0,\ \mathrm{Tr}\, \varrho(x) = 1 \quad (219)$$

$$F(\varrho_i,\varrho_j) = \{\mathrm{Tr}[(\varrho_i^{1/2}\, \varrho_j\, \varrho_i^{1/2})^{1/2}]\}^2 \quad (220)$$

$$D\_B(\varrho_i,\varrho_j) = [2 - 2 \sqrt{F}(\varrho_i,\varrho_j)]^{1/2} \quad (221)$$

$$d\ s^2\_FS = \langle d\psi | d\psi \rangle - |\langle \psi | d\psi \rangle|^2 \quad (222)$$

$$g\_ab(x) = \mathrm{Re}\ \mathrm{Tr}[(\partial_a \varrho(x))(\partial_\beta \varrho(x))] \quad (223)$$

$$F\_ab(x) = 4\ g\_ab(x) \quad (224)$$

$$\Phi\_hyp(r) = {}_2F_1(a,b;c; -\lambda\ r^2) \quad (225)$$

$$\Phi\_CH(r) = \Sigma_{n=0}^{N}\ c_n\ T_n(r)\ \Phi\_hyp(r) \quad (226)$$

$$\Phi\_gauge(r) = \Phi\_CH(r)\ \exp[-\beta\ |\Gamma(r)|] \quad (227)$$

$$A\_a(x) = \mathrm{Im}\ \langle \psi(x) | \partial_a \psi(x) \rangle \quad (228)$$

$$\Omega\_ab(x) = \partial_a A\_b(x) - \partial_\beta A\_a(x) \quad (229)$$

$$E\_geo(\gamma) = \int_0^1 \dot{\gamma}^a(t)\ g\_ab(\gamma(t))\ \dot{\gamma}^b(t)\ dt \quad (230)$$

$$E\_ij = 1\{\bar{K}\_ij \geq \tau\}\ 1\{D\_B(\varrho_i,\varrho_j) \leq \delta\}\ 1\{P\_geo \leq P\_max\} \quad (231)$$

$$A\_ij = \bar{K}\_ij\ E\_ij \quad (232)$$

$$D\_ii = \Sigma_j\ A\_ij \quad (233)$$

$$U_0 = (1 - \alpha)\ D^{-1}\ A + \alpha\ n^{-1}\ 11^T \quad (234)$$

$$U(t) = U_0^t \text{ or } U(t) = \exp[-i\ t\ \tilde{L}] \quad (235)$$

$$S\_path(L) = T^{-1} \int_0^T \Sigma_j{\in}L\ p_j(t)\ dt \quad (236)$$

$$\Phi\_DFT = (E\_KS, \Delta\varepsilon\_HL, \mu, \eta, \omega, q\_res, \mu\_dip, c\_SCF) \quad (237)$$

$$S\_DFT = r\_DFT\ \beta\_DFT^T\ \Phi\_DFT \quad (238)$$

$$E\_QMC = \langle \Psi\_T | H | \Psi\_T \rangle / \langle \Psi\_T | \Psi\_T \rangle \quad (239)$$

$$\sigma^2\_QMC = (M - 1)^{-1}\ \Sigma_m\ [E\_L(R_m) - E\_QMC]^2 \quad (240)$$

$$S\_QMC = -E\_QMC - \lambda\_\sigma\ \sigma\_QMC - \lambda\_CI\ CI\_QMC \quad (241)$$

$$\Pi\_diag = \lambda\_D(1 - r\_DFT) + \lambda\_Q(1 - r\_QMC) + \lambda\_T\ \Delta\_topo + \lambda\_M\ I\_missing \quad (242)$$

$$\Pi\_QG = \lambda\_B\ \Sigma_i D\_B^2(\varrho_i,\varrho_{i+1}) + \lambda\_g\ E\_geo(\gamma) + \lambda\_\Omega \int\gamma\ \|\Omega\_ab\|^2\ dt \quad (243)$$

$$S\_QG^+(P,L) = w_0 S\_\theta + w\_D\ S\_DFT + w\_Q\ S\_QMC + w\_M\ S\_MQW + w\_path\ S\_path - \Pi\_diag - \Pi\_QG \quad (244)$$

$$R^{+}(P,L) = \text{rank\_L}[\text{S\_QG}^{+}(P,L)],\ \Delta R(P,L) = R_0(P,L) - R^{+}(P,L) \tag{245}$$

$$\text{S\_cons}^{+}(L) = \text{Agg\_P}\{\text{S\_QG}^{+}(P,L)\} \tag{246}$$

$$\text{Q\_ready}(L) = 1\{\text{r\_DFT} \geq \text{r\_min}\}\ 1\{\text{r\_QMC} \geq \text{q\_min}\}\ 1\{\text{S\_path} \geq \kappa\text{\_path}\}\ 1\{\Pi\text{\_QG} \leq \pi\text{\_max}\} \tag{247}$$

$$\text{Cert}(L) = (\text{role order, planar feasibility, PSD overlap, Bures bound, curvature control, MQWalk support, DFT/QMC audit}) \tag{248}$$

$$L^{*} = \text{arg max\_L S\_QG}^{+}(P,L) \text{ subject to Q\_ready}(L) = 1 \tag{249}$$

$$\text{Decision}(L) = \text{promote if } \Delta R > 0 \text{ and Q\_ready} = 1\text{; hold if } \Delta R = 0\text{; demote or flag otherwise} \tag{250}$$

$$\text{C\_QG}(L) = [\bar{K}, \text{D\_B, g\_ab, F\_ab, }\Omega\text{\_ab, S\_path, r\_DFT, r\_QMC, S\_QG}^{+}] \tag{251}$$

$$\text{I\_audit}(L) = \{\text{Figures 10-14, Tables 12-15, SI Appendix I Figs. 1-9}\} \tag{252}$$

$$\text{Output} = (\text{leaderboard}, \Delta R, \text{C\_QG, Cert, flags, reproducibility manifest}) \tag{253}$$

Equations (219)-(253) define the expanded quantum-geometric audit layer used to interpret the HMC manifold (Figure 10), the hypergeometric DFT/QMC/MQWalk circuit (Figure 11), the proxy-overlay generalization surface (Figure 12), the diagnostic reporting matrix (Figure 13), the topology-validation graph (Figure 14), and the corresponding reporting tables (Tables 12-15). They also establish a reference-based link between the main text, SI Appendix I figures, and the reproducibility algorithms in SI Appendix II.

3. Results

The Results are organized around the reference-based reporting architecture introduced in Figures 10-14 and Tables 12-15. The main manuscript figures define the quantum-geometric objects, the score-audit circuit, the proxy-overlay generalization surface, the diagnostic matrix, and the MQWalk topology-validation graph. SI Appendix I Figs. 1-9 are cited as supporting visual evidence for the biological, chemical, docking, DFT/QMC, MQWalk, and hotspot contexts, whereas SI Appendix II records the algorithmic pseudocode used to reproduce the audit workflow.

3.1. Overview of Rank Behavior: Auditable Re-Ranking and Near-Tie Resolution

Across the GLYBATOMAQ™ Hyper Mega Core variants, HSX consensus representatives, and family-stratified comparator analyses, the workflow behaved as a late-stage, audit-forward re-ranking layer rather than as a replacement docking leaderboard. This point is central to the interpretation of the results and formally resolves the methodological possibility that the method might otherwise appear as an overly complex scoring construction. The numerical outcomes are not presented as claims of binding affinity, biological efficacy, or clinical performance. Instead, they are interpreted as a structured mathematical-chemistry ranking experiment in which each candidate's movement is examined through docking position, positive-semidefinite operator overlap, DFT descriptor behavior, QMC uncertainty, MQWalk topology consistency, quantum-geometric smoothness, curvature regularity, and diagnostic penalties. The geometric basis for this interpretation is summarized in Figure 10, which shows how HMC role grammar, PSD-kernel connectability, Bures/Fubini-Study-style distance control, MQWalk support, and MBHA projection convert candidate generation into a constrained two-dimensional quantum-geometric pharmacophore path rather than an arbitrary scaffold-merging operation. These results are read through the quantum-geometric reporting schema of Figure 13 and Tables 12-15, while Figures 10-12 and Figure 14 provide the manifold, circuit, proxy-overlay, and topology-validation views used to interpret rank movement.

Docking remained the protocol-fixed upstream signal. AutoDock Vina-derived and DockThor-derived values were treated as rank-bearing oracle outputs under fixed computational settings, consistent with the intended comparative use of docking scores [6,13] and the broader cautions on

doi:10.20944/preprints202605.2096.v1

docking-score interpretation emphasized by Kitchen et al. [9], Warren et al. [10], Pagadala et al. [11], Guedes et al. [12], PDBbind benchmarking [14], Huang et al. [15], and DUD-E [16]. Thus, the primary result was not a single final scalar but a rank-audit object,

$$R^{+}(P,L) = \mathrm{rank}_{l}[S\_QG^{+}(P,L)],\ \Delta R = R_{0}(P,L) - R^{+}(P,L) \tag{254}$$

where the explicit rank shift was

$$D_{\beta\mathrm{est}}(L) = \mathrm{min}P \in \mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\mathrm{min}}L' D_{\beta\mathrm{est}}(L') \tag{255}$$

Positive ΔR indicates promotion after refinement, negative ΔR indicates demotion, and zero indicates rank stability. This is the central observable of the Results section. The same rank-first philosophy is shown in SI Appendix I Figure 1A, where docking/oracle scoring is stabilized by an operator-kernel/pattern-recognition layer, followed by branched DFT, QMC, and MQWalk refinement before reliability-aware fusion and report/export.

The clearest result was observed in **near-tie regimes**. These are regions where docking scores, operator-overlap similarities, or component-energy channels are closely spaced, so that a purely scalar leaderboard would be unstable or chemically underexplained. In such cases, the GLYBATOMAQ™ refinement layer did not replace docking. It resolved ambiguity by asking whether a promotion was supported by deterministic electronic structure, stochastic energetic stability, and graph-topology consistency.

The rank-audit logic can be summarized as

$$\mathrm{SDFT}(P,L) = a^{T}\Phi\mathrm{DFT}(P,L) + b,\ \mathrm{CDFT} = \mathrm{rDFT}\ \mathrm{SDFT} \tag{256}$$

This behavior is visible across SI Appendix I Figure 5, where DFT-branch descriptors and the fusion circuit are placed alongside docking-energy baselines, and SI Appendix I Figure 6, where topology validation, hotspot landscapes, and residue-level fingerprints are integrated into a single audit chain. The SI Appendix explicitly states that DFT/QMC/MQWalk are used for auditable re-ranking, uncertainty handling, and topology validation rather than for absolute binding claims.

In the integrated figure schema, Figure 13 acts as the results-level reporting matrix for these rank movements: PSD overlap, Bures control, Fubini-Study geometry, quantum Fisher sensitivity, Berry-type gauge consistency, DFT reliability, QMC uncertainty, MQWalk support, and fusion readiness are treated as normalized evidence channels rather than as isolated metrics. Figure 14 then gives the corresponding topology interpretation by showing how a promotion is supported only when propagated mass remains in compatible operator-overlap neighborhoods. These main-text figures are cited together with SI Appendix I Figs. 1, 3, 5, and 6 because the SI panels provide the circuit, transport, DFT-branch, docking-energy, hotspot, and residue-fingerprint evidence underlying the summarized reporting channels.

From a Pattern Recognition standpoint, the results are naturally interpreted as ranking on structured chemical data rather than scalar regression. This is aligned with ranking-based learning [56], local-kernel graph learning [57], informative graph selection [58], dynamic graph representation [59], adaptive graph propagation [60], and molecular graph contrastive learning [61]. The chemical representation layer remained chemistry-safe, using MMFF94-compatible geometry handling [17], RDKit descriptors [18], Morgan-style encodings [19], ECFP fingerprints [20], SMILES [21], SELFIES [22], and structured molecular-learning compatibility [23-32].

In practical terms, the principal result of the re-ranking layer was not that a single candidate “won.” Rather, the Results section shows that several candidates occupied rank-sensitive regions where additional evidence was necessary. GLYBATOMAQ-HSX-CONS2 and GLYBATOMAQ-HSX-cons4 formed the most important near-tie pair in the AutoDock Vina-derived best-pocket analysis, while GLYBATOMAQ-HSX-cons2, Vorasidenib, GLYBATOMAQ-HSX-cons1, and GLYBATOMAQ-HSX-cons4 formed a closely spaced group in the DockThor 3Q2R analysis. This justified the use of branch-level audit rather than isolated score interpretation.

3.2. Docking-Defined Frontier and Numerical Rank Baselines

The first numerical result is the docking-defined frontier. In the GLYBATOMAQ™ framework, this frontier is not interpreted as a binding-affinity table. It is a controlled computational ranking baseline from which refinement begins.

For the AutoDock Vina pocket-resolved analysis, the best-pocket value was defined as

$$\mathrm{SQMC}(P,L) = -\bar{E}\mathrm{QMC}(P,L) - \lambda\sigma\ \sigma\mathrm{QMC}(P,L) - \lambda\mathrm{CI}\ \mathrm{CI\_QMC}(P,L) \tag{257}$$

where $E(L,p)$ is the docking score of ligand L in pocket p. More negative values indicate a more favorable docking score under the same protocol. The gap-to-best statistic was

$$\mathrm{SMQW}(P,L) = \lambda_{\mathrm{path}}\ S_{\mathrm{path}}(P,L) - \lambda^{d}{}_{\mathrm{isc}}\ \Delta_{\mathrm{topo}}(P,L) \tag{258}$$

This statistic is important because it makes near ties explicit. A candidate separated from the best ligand by $\Delta_{\mathrm{best}} = 0.1$ should not be treated as decisively inferior, whereas a candidate separated by $\Delta_{\mathrm{best}} > 2$ belongs to a different docking tier under the fixed protocol.

Table 1. AutoDock Vina best-pocket leaderboard and near-tie gaps.

Rank	Candidate	Structural context	Best pocket	E_{min} kcal/mol	Δ_{best} kcal/mol	Rank interpretation
1	GLYBATOMAQ-HSX-CONS2	2GLI	C2/C5 tie	-10.4	0.0	Best docking-frontier candidate
2	GLYBATOMAQ-HSX-cons4	3Q2R	C1	-10.3	0.1	Near-tie with top candidate
3	Bafetinib	2GLI	C1	-8.5	1.9	Start of second tier
4	GLYBATOMAQ-HSX-cons1	2GLI	C2	-7.7	2.7	Mid-ranked HSX candidate
5	Crenolanib	2GLI	C2	-7.6	2.8	Comparator near cons1
6	Tovorafenib	2GLI	C2	-7.5	2.9	Comparator near cons1/Crenolanib
7	Galunisertib	—	C1	-7.2	3.2	Mid-tier comparator
8	Trametinib	—	C4	-7.1	3.3	Mid-tier comparator
9	Olutasidenib	—	C1	-6.9	3.5	Lower-mid comparator
10	AZD8055	—	C1	-6.8	3.6	Lower-mid comparator
11	Paxalisib	—	C2	-6.7	3.7	Lower-mid comparator
12	Vorasidenib	—	C1	-6.5	3.9	Lower docking-frontier comparator
13	Temozolomide	—	C1	-5.4	5.0	Weak docking-frontier comparator
14	Indoximod	—	C2	-5.3	5.1	Weakest listed comparator

The key result from Table 1 is that the top two HSX candidates are separated by only 0.1 kcal/mol under the best-pocket Vina readout. This is the most important near-tie in the AutoDock Vina analysis. GLYBATOMAQ-HSX-CONS2 is ranked first, but GLYBATOMAQ-HSX-cons4 is close enough that a responsible interpretation should not treat the ordering as chemically decisive without additional audit. This is precisely the region where DFT, QMC, and MQWalk are intended to add mathematical-chemistry value. SI Appendix I Figure 7 and SI Appendix I Figure 8 explicitly frame these values as docking-oracle ranking artifacts and use the gap-to-best statistic to avoid overinterpretation of small score differences.

The second numerical docking baseline came from the comparative docking-energy overview in SI Appendix I Figure 5B. This panel compared four GLYBATOMAQ HSX consensus representatives against selected FDA-approved cancer-drug comparators under the reported docking-energy overview.

Table 2. Comparative docking-energy overview: HSX consensus representatives versus FDA comparators.

Rank	Candidate	Reported docking-energy overview value	Interpretation within GLYBATOMAQ™
1	HSX-cons2	-125.2	Strongest HSX docking-frontier value in this overview
2	HSX-cons1	-113.9	Second HSX candidate; remains clearly separated from comparators
3	HSX-cons3	-95.6	Third HSX candidate
4	HSX-cons4	-92.0	Fourth HSX candidate
5	Vemurafenib	-72.1	Strongest listed FDA comparator in this overview
6	Lorlatinib	-65.5	Comparator
7	Ribociclib	-64.4	Comparator
8	Osimertinib	-62.3	Comparator
9	Sorafenib	-60.2	Comparator
10	Brigatinib	-57.9	Comparator

These values show that, under the reported docking-energy overview, the four HSX consensus representatives occupy the top docking tier relative to the listed comparators. However, the interpretation remains conservative. The values define a pre-refinement ordering under a fixed computational protocol. They do not imply experimental affinity or efficacy. SI Appendix I Figure 5B explicitly states that the comparative docking-energy overview is used as a rank-bearing baseline and that DFT/QMC/MQWalk act as auditable re-ranking, uncertainty-handling, and topology-validation branches.

Figure 12 converts these chemical-energy comparisons into a proxy-overlay surface by placing HSX docking-energy proxies, MMP3 inhibitor ranking anchors, and candidate-level quantum-geometry supports on the same normalized interpretive plane. This design allows the Results section to compare chemical and biological evidence in a reference-based style while preserving the conservative boundary that docking, MD, DFT, QMC, and MQWalk outputs remain computational prioritization signals.

A third docking baseline came from the DockThor 3Q2R panel. DockThor affinity was again treated as a rank-bearing docking-oracle score, with more negative values ranked as more favorable.

Table 3. DockThor 3Q2R affinity leaderboard.

Rank	Candidate	DockThor affinity	Interpretation
1	GLYBATOMAQ-HSX-cons2	-9.100	Best DockThor affinity in this panel
2	Vorasidenib	-8.350	Strong comparator, close to HSX-cons1/cons4
3	GLYBATOMAQ-HSX-cons1	-8.295	HSX candidate in near-tie region
4	GLYBATOMAQ-HSX-cons4	-8.289	HSX candidate nearly tied with cons1
5	Carmustine	-7.976	Comparator
6	Temozolomide	-7.866	Comparator
7	GLYBATOMAQ-HSZ-cons3	-7.793	HSZ/HSX-family candidate in lower part of this DockThor panel

The DockThor panel shows a different but still rank-informative ordering. GLYBATOMAQ-HSX-cons2 remained the top candidate in this 3Q2R DockThor readout, whereas Vorasidenib, GLYBATOMAQ-HSX-cons1, and GLYBATOMAQ-HSX-cons4 formed a closely spaced group. This confirms the need for pocket-aware and engine-aware interpretation. A single docking engine or pocket should not be overread; instead, differences across Vina, DockThor, and pocket contexts become inputs to the audit layer. SI Appendix I Figure 9 reports the DockThor affinity panel, component-energy vector, reproduced numeric table, and screenshot audit trail, preserving the manuscript's rank-centric stance.

3.3. Near-Tie Regimes as the Main Site of Mathematical-Chemistry Gain

The numerical results identify two types of near-tie regimes.

The first type is a **score-near-tie**, where two or more candidates have very similar docking values. The HSX-CONS2 versus HSX-cons4 AutoDock Vina comparison is the clearest example:

$$S\_QG^{+} = w_0 S\theta + wD\ CDFT + wQ\ CQMC + wM\ SMQW - \Pi\_diag - \Pi_{curv} \tag{259}$$

A 0.1 kcal/mol difference in docking score is too small to support a strong chemical conclusion on its own. In GLYBATOMAQ™, such a difference is treated as a trigger for audit, not as a final result.

The second type is a **component-near-tie**, where the final affinity or total score is close, but the underlying energy components differ. SI Appendix I Figure 9 gives such an example: GLYBATOMAQ-HSX-cons1 and Vorasidenib are close in DockThor affinity, but their reported energy-component patterns differ. GLYBATOMAQ-HSX-cons1 combines DockThor affinity of -8.295 with $E_{\mathrm{vdW}} = -10.802$ and a strongly negative electrostatic component $E_{\mathrm{elec}} = -24.888$, whereas Vorasidenib has affinity -8.350 with $E_{\mathrm{total}} = -0.825$, $E_{\mathrm{vdW}} = -15.779$, and $E_{\mathrm{elec}} = -2.284$. These differences mean that similar affinity-scale outputs can arise from different energetic patterns.

This distinction is important because it gives our Results section a clear mathematical-chemistry message: **rank similarity does not imply mechanistic similarity**. Two candidates may occupy adjacent ranks while differing in electronic, topological, or residue-contact structure. This is why GLYBATOMAQ™ reports branch contributions, topology flags, and residue-level fingerprints rather than only final ranks.

3.4. DFT Refinement: Deterministic Electronic Separation of Ambiguous Candidates

The DFT branch contributed deterministic electronic-structure evidence to the rank audit. It was not used to claim absolute binding energies. Instead, it supplied descriptor-level separation among candidates whose docking and operator-overlap signals were insufficiently discriminative.

The DFT branch follows the Hohenberg–Kohn formalism [44], the Kohn–Sham equations [45], and standard DFT treatments [46,47]. For each logged geometry $\chi^{\star}(P, L)$, the formal basis is

$$\mathcal{O}@k = |\mathrm{Top}_k(R_0) \cap \mathrm{Top}_k(R^{+})| / k \tag{260}$$

with

$$\varrho S = \mathrm{corrS}_{\mathrm{pearman}}(R_0, R^{+}) \tag{261}$$

and stationarity condition

$$R^{+}(P,L) = \mathrm{rank}_l[S\_QG^{+}(P,L)],\ \Delta R = R_0(P,L) - R^{+}(P,L) \tag{262}$$

Operationally, the DFT branch generated a descriptor vector,

$$D_{\beta\mathrm{est}}(L) = \mathrm{min}P\in\mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\mathrm{min}}L'D_{\beta\mathrm{est}}(L') \tag{263}$$

which was mapped to a branch score,

$$SDFT(P,L) = a^{\mathrm{T}}\Phi DFT(P,L) + b,\ CDFT = rDFT\ SDFT \tag{264}$$

Reliability was carried explicitly through

$$SQMC(P,L) = -\bar{E}QMC(P,L) - \lambda\sigma\ \sigma QMC(P,L) - \lambda CI\ CI\_QMC(P,L) \tag{265}$$

The reliability gate is essential. Without it, an unstable DFT calculation could create a misleading rank promotion. In the GLYBATOMAQ™ results, SCF failures, geometry-sensitive outliers, descriptor-domain violations, or incomplete DFT outputs were not silently discarded. They were carried as reliability reductions or diagnostic penalties:

$$SMQW(P,L) = \lambda_{\mathrm{path}}\ S_{\mathrm{path}}(P,L) - \lambda^{d}{}_{\mathrm{isc}}\ \Delta_{\mathrm{topo}}(P,L) \tag{266}$$

Thus, DFT evidence could strengthen a candidate only when the electronic calculation was stable enough to support interpretation.

SI Appendix I Figure 5A visualizes the DFT branch using descriptor histograms and a descriptor-plane scatter plot. The figure caption describes these histograms as empirical distributions of DFT-derived electronic descriptors over the shortlisted set and the scatter plot as a near-tie discrimination view. This is precisely where the DFT branch adds research-level mathematical-chemistry value: it transforms ambiguous rank differences into interpretable electronic descriptor differences, while preserving reliability flags.

The DFT result can therefore be stated as follows: DFT did not function as an independent replacement score; it functioned as a **deterministic near-tie separator**. In candidates whose docking values were close, DFT descriptors provided a second axis of evidence. However, only reliable DFT contributions were allowed to affect the fused rank.

3.5. QMC Refinement: Uncertainty-Aware Energetic Auditing Rather than Screening-Scale Rescoring

The QMC branch had a narrower and more conservative role than DFT. It was applied only to strict top subsets or rank-sensitive regions, where stochastic energetic uncertainty could materially alter interpretation. This role is consistent with QMC methodology [48,49] and with the overall GLYBATOMAQ™ boundary: QMC is an audit layer, not a screening-scale rescoring replacement.

The QMC branch was defined through the energy expectation,

$$S\_QG^{+} = w_0 S\theta + wD\ CDFT + wQ\ CQMC + wM\ SMQW - \Pi\_diag - \Pi_{curv} \tag{267}$$

and a diffusion-like refinement view,

$$\mathcal{O}@k = |\mathrm{Top}_k(R_0) \cap \mathrm{Top}_k(R^{+})| \,/\, k \tag{268}$$

The practical QMC output was encoded as

$$\varrho S = \mathrm{corrS}_{\mathrm{pearman}}(R_0, R^{+}) \tag{269}$$

with conservative score

$$R^{+}(P,L) = \mathrm{rank}_l[S\_QG^{+}(P,L)],\ \Delta R = R_0(P,L) - R^{+}(P,L) \tag{270}$$

This expression made the QMC result uncertainty-aware by construction. A favorable $\hat{E}_{\mathrm{QMC}}$ was not sufficient for rank support if σ_{QMC} or CIwidth was large. In other words, QMC separated candidates with reproducible energetic support from those whose apparent gains were dominated by stochastic noise.

This is the correct level of claim for QMC in the present manuscript. The branch does not assert exhaustive quantum-level binding energetics. It provides selective uncertainty auditing. The SI Appendix describes the QMC branch as an uncertainty-explicit audit layer that contributes a conservative uncertainty-penalized score and a reliability gate.

The QMC result may therefore be summarized as:

$$D_{\beta\mathrm{est}}(L) = \mathrm{min}P \in \mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\mathrm{min}}L' D_{\beta\mathrm{est}}(L') \tag{271}$$

Only candidates satisfying

$$SDFT(P,L) = a^{T}\Phi DFT(P,L) + b,\ CDFT = rDFT\ SDFT \tag{272}$$

received positive QMC support in the final rank audit.

QMC is not introduced as an ornate computational label. It performs a necessary role: it prevents the algorithm from treating uncertain energetic estimates as if they were precise.

3.6. Reliability-Aware DFT/QMC/MQWalk Fusion

The final fused score was

$$SQMC(P,L) = -\bar{E}QMC(P,L) - \lambda\sigma\ \sigma QMC(P,L) - \lambda CI\ CI\_QMC(P,L) \tag{273}$$

This equation is not merely a scoring formula. It is the mathematical structure that makes the Results interpretable. Each term has a nonredundant role:

$$SMQW(P,L) = \lambda_{\mathrm{path}}\ S_{\mathrm{path}}(P,L) - \lambda^{d}_{\mathrm{isc}}\ \Delta_{\mathrm{topo}}(P,L) \tag{274}$$

preserves the upstream ranking frontier;

$$S\_QG^{+} = w_0 S\theta + wD\ CDFT + wQ\ CQMC + wM\ SMQW - \Pi\_diag - \Pi_{curv} \tag{275}$$

adds deterministic electronic support only when DFT reliability is acceptable;

$$\mathcal{O}@k = |\mathrm{Top}_k(R_0) \cap \mathrm{Top}_k(R^+)| / k \tag{276}$$

adds uncertainty-controlled energetic support;

$$\varrho S = \mathrm{corrS}_{\mathrm{pearman}}(R_0, R^+) \tag{277}$$

adds topology-consistency support;

$$R^+(P,L) = \mathrm{rank}_l[S\_QG^+ (P,L)],\ \Delta R = R_0(P,L) - R^+(P,L) \tag{278}$$

penalizes unreliable calculations, topology discordance, and missing provenance.

The diagnostic penalty can be written as

$$D_{\beta\mathrm{est}}(L) = \min P \in \mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\min} L' D_{\beta\mathrm{est}}(L') \tag{279}$$

Thus, the final score is not a black-box ensemble. It is a reliability-aware evidence-integration rule. SI Appendix I Figure 5B shows this fused sequence as Scale → Expand → Aggregate → Score → Compare → Refine → Validate → Re-rank, with docking feeding a shortlist and branching into DFT, QMC, and MQWalk before final rank audit.

The main result of the fusion layer was interpretability. A candidate could be promoted because DFT descriptors and QMC uncertainty both supported the movement and MQWalk did not flag topology discordance. Conversely, a candidate could be demoted or held stable if its apparent gain depended on one weak branch, unstable uncertainty, or graph-neighborhood disagreement.

3.7. Chemistry-Safe HMC/HSX Consensus Behavior

The hierarchical GLYBATOMAQ™ Hyper Mega Core behaved as a chemistry-safe consensus abstraction rather than an atom-exact mega-scaffold. This point is important because the HMC could otherwise be misread as a chemically fused structure. In the Results interpretation, the HMC is a numerical consensus geometry for descriptor-space comparability, docking-oriented pattern recognition, and audit-forward re-ranking.

The HMC map encoded pharmacophore-role modules:

$$\mathrm{SDFT}(P,L) = a^T \Phi \mathrm{DFT}(P,L) + b,\ \mathrm{CDFT} = \mathrm{rDFT\ SDFT} \tag{280}$$

Edges represented within-family descriptor compatibility, while thin teleport/refinement edges connected comparator anchors in kernel space. The SI Appendix explicitly states that the HMC is a numerical, chemistry-safe consensus geometry, not an atom-exact “mega-scaffold.”

This interpretation was reinforced by the HSX dashboard. SI Appendix I Figure 3 displayed HSX-cons1–4 as fixed reference architectures and separated these anchors from reporting plots used to interpret rank behavior, refinement reliability, and topology consistency. The dashboard also included interaction-envelope heatmaps, MQWalk-style propagation traces, and uncertainty/dispersion plots, all framed as audit surfaces rather than biological proof.

The chemistry-safe consensus behavior is consistent with RDKit [18], SELFIES [22], grammar VAE [23], junction-tree VAE [24], graph-policy learning [25], recurrent design [28], latent molecular design [29], message passing [30], graph convolution [31], and attention-based representation learning [32]. It is also compatible with Pattern Recognition interest in structured graph representations and molecular graph learning [59-61].

The result is that GLYBATOMAQ™ preserves family-level comparability without forcing heterogeneous chemotypes into a misleading atom-level consensus. This is a methodological strength, not a complication: it prevents invalid chemical fusion while allowing structured comparison.

3.8. MQWalk Topology Validation and Operator-Overlap Consistency

MQWalk functioned as a topology-consistency guardrail. It projected DFT- and QMC-driven rank movement back onto the normalized operator-overlap graph and tested whether promotions were neighborhood-consistent.

Pocket–ligand comparability was anchored by PSD operators,

$$\mathrm{SQMC}(P,L) = -\bar{\mathrm{E}}\mathrm{QMC}(P,L) - \lambda\sigma\ \sigma\mathrm{QMC}(P,L) - \lambda\mathrm{CI}\ \mathrm{CI\_QMC}(P,L) \tag{281}$$

typically constructed as

$$\mathrm{SMQW}(P,L) = \lambda_{\mathrm{path}}\ S_{\mathrm{path}}(P,L) - \lambda^{d}{}_{\mathrm{isc}}\ \Delta_{\mathrm{topo}}(P,L) \tag{282}$$

The trace overlap was

$$\mathrm{S\_QG}^{+} = w_{0}S\theta + w\mathrm{D\ CDFT} + w\mathrm{Q\ CQMC} + w\mathrm{M\ SMQW} - \Pi\mathrm{\_diag} - \Pi_{\mathrm{curv}} \tag{283}$$

and the normalized trace overlap was

$$\mathcal{O}@k = |\mathrm{Top}_{k}(R_{0}) \cap \mathrm{Top}_{k}(R^{+})| / k \tag{284}$$

This operator-overlap geometry is consistent with kernel learning [33], Gaussian-process similarity concepts [35], and positive-definite matrix theory [36]. It is also the mathematical foundation of the MQWalk branch.

From $\bar{K}$, the bipartite adjacency was

$$\varrho S = \mathrm{corr}S_{\mathrm{pearman}}(R_{0}, R^{+}) \tag{285}$$

The degree-normalized adjacency was

$$R^{+}(P,L) = \mathrm{rank}_{l}[\mathrm{S\_QG}^{+}(P,L)],\ \Delta R = R_{0}(P,L) - R^{+}(P,L) \tag{286}$$

Teleport mixing gave

$$D_{\beta\mathrm{est}}(L) = \min P \in \mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\min}L'D_{\beta\mathrm{est}}(L') \tag{287}$$

The Avogadro-conditioned Laplacian was

$$\mathrm{SDFT}(P,L) = a^{T}\Phi\mathrm{DFT}(P,L) + b,\ \mathrm{CDFT} = r\mathrm{DFT\ SDFT} \tag{288}$$

where

$$\mathrm{SQMC}(P,L) = -\bar{\mathrm{E}}\mathrm{QMC}(P,L) - \lambda\sigma\ \sigma\mathrm{QMC}(P,L) - \lambda\mathrm{CI}\ \mathrm{CI\_QMC}(P,L) \tag{289}$$

and N_A was used only as a CODATA-based numerical scale anchor [43].

The continuous-time walk operator was

$$\mathrm{SMQW}(P,L) = \lambda_{\mathrm{path}}\ S_{\mathrm{path}}(P,L) - \lambda^{d}{}_{\mathrm{isc}}\ \Delta_{\mathrm{topo}}(P,L) \tag{290}$$

and node mass was

$$\mathrm{S\_QG}^{+} = w_{0}S\theta + w\mathrm{D\ CDFT} + w\mathrm{Q\ CQMC} + w\mathrm{M\ SMQW} - \Pi\mathrm{\_diag} - \Pi_{\mathrm{curv}} \tag{291}$$

The MQWalk score was

$$\mathcal{O}@k = |\mathrm{Top}_{k}(R_{0}) \cap \mathrm{Top}_{k}(R^{+})| / k \tag{292}$$

The key result was that MQWalk did not behave as a new physical score source. It functioned as a mathematical constraint on whether a post-refinement promotion was coherent with the learned overlap topology. SI Appendix I Figure 1B–C and SI Appendix I Figure 6 present topology/transport diagnostics, heatmaps, bin-resolved CTQW-style summaries, and MQWalk/teleport validation panels as algorithmic consistency checks, not as claims of biomolecular quantum transport.

This interpretation is consistent with quantum-information formalism [50] and with cautions regarding quantum-biological overinterpretation [51,52]. The Results therefore use MQWalk in the restricted sense needed for mathematical chemistry: graph-propagation validation of rank movement.

3.9. Residue-Resolved Energetic Fingerprints and Hotspot-Supported Interpretation

The topology results became more informative when paired with residue-level energetic fingerprints and hotspot landscapes. SI Appendix I Figure 6 stacked global hotspot/feature landscapes, MQWalk-style topology validation, and residue-resolved energetic fingerprints into a single auditable evidence chain. The figure explicitly framed docking as a fixed-protocol oracle, PSD operator kernels as the stabilizing machinery, MQWalk as topology validation, and residue/contact decomposition as the bottom-up energetic audit trail.

The residue-level interface score was written as

$$\varrho S = \mathrm{corrS}_{\mathrm{pearman}}(R_0, R^+) \tag{293}$$

where R is the residue/contact vocabulary. Each ligand can therefore be represented as a sparse energetic fingerprint,

$$R^+(P,L) = \mathrm{rank}_l[S\_QG^+ (P,L)],\ \Delta R = R_0(P,L) - R^+(P,L) \tag{294}$$

The nonzero support

$$D_{\beta\mathrm{est}}(L) = \min P\in\mathcal{P}\ D(P,L),\ \mathrm{gap}(L) = D_{\beta\mathrm{est}}(L) - {}_{\min}L'D_{\beta\mathrm{est}}(L') \tag{295}$$

defines the ligand's hotspot signature.

Table 4. Residue-resolved energetic fingerprint: model1_2gli–GLYBATOMAQ-HSX-cons1.

Candidate	Total energy	Strong stabilizing contributors	Counteracting positive contributor	Interpretation
model1_2gli–GLYBATOMAQ-HSX-cons1	-100.1	$\Delta E_{\mathrm{V-M-GLN154}} = -13.8519$; $\Delta E_{\mathrm{V-S-ARG163}} = -10.0914$; $\Delta E_{\mathrm{V-M-GLU119}} = -9.75054$; $\Delta E_{\mathrm{V-S-TRP108}} = -8.23134$	$\Delta E_{\mathrm{V-S-HIS14}} = +6.76114$	Strong negative contributors are balanced by at least one destabilizing/local positive term

Table 5. Residue-resolved energetic fingerprint: model1_2gli–GLYBATOMAQ-HSX-cons2.

Candidate	Total energy	Strong stabilizing contributors	Positive offsets	Interpretation
model1_2gli–GLYBATOMAQ-HSX-cons2	-97.5	$\Delta E_{\mathrm{V-M-HIS164}} = -14.0114$; $\Delta E_{\mathrm{V-M-SER226}} = -11.1797$; $\Delta E_{\mathrm{V-S-ARG178}} = -8.36414$; $\Delta E_{\mathrm{V-S-TYR181}} = -6.23947$; $\Delta E_{\mathrm{V-M-CYS177}} = -5.61929$; $\Delta E_{\mathrm{V-S-GLU185}} = -5.46789$	$\Delta E_{\mathrm{V-S-GLU119}} = +4.99914$; $\Delta E_{\mathrm{V-S-TRP141}} = +1.65004$	Similar total-energy scale to cons1 but different stabilizing wiring pattern

The result is that candidates with comparable total energetic values can have substantially different residue-level "wiring diagrams." This is important because scalar totals alone can hide mechanistic differences. For example, cons1 and cons2 are both supported by strong negative residue terms, but the dominant residues differ: cons1 is characterized by GLN154, ARG163, GLU119, and TRP108 contributions, whereas cons2 is characterized by HIS164, SER226, ARG178, TYR181, CYS177, and GLU185 contributions.

This gives the Results section a clear research-level mathematical-chemistry interpretation:

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \geq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{296}$$

Similar total scores can arise from different sparse energetic supports. This supports the manuscript's claim that rank movement should be interpreted through residue-level fingerprints, hotspot signatures, and topology validation rather than opaque single-number claims.

3.10. Interaction Persistence, SI-Derived DFT/QMC/MQWalk Audit, and Biology-Facing Stability Diagnostics

The residue/contact decomposition gives the Results section a clear research-level mathematical-chemistry interpretation. In particular,

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{297}$$

Equation (296) states that similar total interface or docking-derived scores do not necessarily imply similar residue-level energetic fingerprints. Two ligands may occupy a similar scalar-energy range while being supported by different sparse stabilizing and destabilizing residue/contact contributions. This distinction is important in docking-led virtual screening because docking scores are best interpreted as comparative, protocol-dependent ranking signals rather than direct experimental binding free energies [6–16]. It is also consistent with the operator-kernel interpretation

doi:10.20944/preprints202605.2096.v1

of GLYBATOMAQ™, in which ligand–pocket comparison is not reduced to a single scalar value but is embedded in a positive-semidefinite similarity geometry [33,35,36].

Accordingly, the present Results interpretation treats rank movement as a structured mathematical-chemistry audit. Residue-level fingerprints, hotspot signatures, interaction persistence, DFT proxy descriptor support, QMC-style uncertainty penalties, and MQWalk topology validation are used to determine whether a candidate's promotion or demotion is explainable. This approach follows the conservative scope of the manuscript: docking provides a fixed-protocol frontier, while the post-docking refinement layers test whether movement within that frontier is electronically plausible, uncertainty-aware, and topologically coherent [9–16,44–52].

The structural basis of the GLIPR1 analysis is anchored to the Protein Data Bank [1], the human GLIPR1 structural study of Asojo et al. [2], and the GLIPR1-related entries 3Q2U and 3Q2R [3,4]. Pocket localization and protein-side interpretation are supported by the Fpocket framework [5]. Within this context, the HSX/HMC representation is treated as a chemistry-safe, rank-comparison geometry rather than as an atom-exact synthetic scaffold. Molecular representation and preprocessing are compatible with MMFF94, RDKit, Morgan/ECFP fingerprints, SMILES, SELFIES, graph representations, and modern molecular learning approaches [17–32].

Using the SI Appendix I numerical values as fixed rank-bearing inputs, the four HSX consensus candidates were converted into a reproducible DFT/QMC/MQWalk branch-score audit. The SI Figure 5B docking-energy overview supplied the main energy-support axis, while the AutoDock Vina best-pocket values from SI Figures 7–8 supplied an independent docking-oracle support axis where available. DockThor values from SI Figure 9 were used only where visible and explicitly reported. Missing values were not invented; when required for a normalized support field, they were imputed from the SI Figure 5B overview and treated as flagged SI-derived values rather than as new calculations.

The resulting SI-derived branch-audit summary was:

Table 6. SI-derived DFT/QMC/MQWalk branch-score audit for HSX-cons1–4.

Candidate	Base support	DFT-style support	QMC reliability	QMC support	MQWalk support	Diagnostic penalty	Fused score	Pre-rank	Post-rank	Rank shift	Interpretation
GLYBATOMAQ-HSX-cons2	1.000	1.000	0.980	0.980	0.922	0.634	0.064	0.852	1	1	0
GLYBATOMAQ-HSX-cons1	0.462	0.589	0.934	0.461	0.833	0.917	0.029	0.503	2	2	0
GLYBATOMAQ-HSX-cons3	0.108	0.305	0.924	0.115	0.778	0.908	0.032	0.235	4	3	+1
GLYBATOMAQ-HSX-cons4	0.282	0.013	0.882	0.126	0.780	0.876	0.034	0.224	3	4	−1

These values show that HSX-cons2 remained the leading candidate after reliability-aware fusion, with . HSX-cons1 remained second, with . The lower-ranked pair illustrates the value of the audit framework: HSX-cons3 moved from rank 4 to rank 3, whereas HSX-cons4 moved from rank 3 to rank 4. This movement occurred because the fused score did not simply reproduce one docking table; it integrated the SI-derived base frontier, DFT proxy descriptor support, QMC-style conservative support, MQWalk topology support, and diagnostic penalties. Thus, the numerical audit operationalizes the manuscript's central claim that rank movement is more informative than an isolated score.

The DFT proxy branch is grounded conceptually in the Hohenberg–Kohn and Kohn–Sham formulation of density-functional theory [44,45], with interpretation consistent with standard DFT treatments [46,47]. In the present SI-derived audit, however, the DFT branch should be read as a deterministic descriptor-support calculation derived from SI-reported score axes, not as a newly executed electronic-structure calculation. Its purpose is to mimic the manuscript-defined role of DFT as a near-tie discriminator: it adds deterministic electronic-style support only when the score landscape provides consistent evidence. This is why HSX-cons2 retained maximal DFT proxy support, whereas HSX-cons4, despite strong Vina near-tie behavior, was weaker on the SI Figure 5B total-energy-proxy axis.

The QMC-style branch follows the uncertainty-aware logic of variational and diffusion Monte Carlo methodology [48,49]. In the present SI-derived audit, QMC support was not treated as a raw energy claim. Instead, the conservative support score penalized uncertainty by subtracting sigma and confidence-width terms from the expected support index. This reflects the mathematical principle that a candidate should not be promoted merely because one favorable score is observed if the supporting evidence is uncertain. In this sense, the QMC-style branch is an uncertainty audit rather than a replacement for experimental binding measurements or full electronic-structure simulation.

The MQWalk branch is based on quantum-walk-style graph propagation notation [50], but its interpretation remains strictly algorithmic. This restriction is important because quantum terminology in biological settings requires caution [51,52]. In GLYBATOMAQ™, MQWalk does not claim physical quantum transport in GLIPR1 or in ligand recognition. It tests whether post-refinement candidate movement remains compatible with the normalized operator-overlap topology derived from the ligand–pocket similarity graph. This is directly aligned with modern graph-learning and pattern-recognition approaches that emphasize ranking, local kernels, graph propagation, dynamic graph representation, and molecular graph contrastive learning [56–61].

The teleport-mixed MQWalk transition matrix used for the HSX topology audit was:

Table 7. Teleport-mixed MQWalk transition matrix for HSX-cons1–4.

From / to	HSX-cons2	HSX-cons1	HSX-cons3	HSX-cons4
H2.	0.037	0.407	0.265	0.291
HSX-cons1	0.287	0.037	0.385	0.290
HSX-cons3	0.194	0.393	0.037	0.376
HSX-cons4	0.230	0.322	0.410	0.037

The corresponding topology outputs were:

Table 8. MQWalk topology diagnostics for HSX-cons1–4.

Candidate	Terminal node mass at HSX node	Stationary centrality	MQWalk topology score	Topology-discordance flag
GLYBATOMAQ-HSX-cons1	0.250	0.279	0.917	0.333
GLYBATOMAQ-HSX-cons3	0.250	0.273	0.908	0.333
GLYBATOMAQ-HSX-cons4	0.250	0.250	0.876	0.333
GLYBATOMAQ-HSX-cons2	0.250	0.198	0.634	1.000

These results show that MQWalk did not simply reward the already top-ranked candidate. Instead, it supplied a topology-consistency diagnostic. HSX-cons2 remained first because its base support, DFT proxy support, and QMC-style support were dominant, even though its topology-discordance term was higher. HSX-cons1, HSX-cons3, and HSX-cons4 had stronger topology support, but their final positions depended on the full reliability-aware balance of branch evidence and penalties.

The biology-facing panel, SI Appendix I Figure 4, extended the same audit-first logic to interaction persistence, MBHA-guided search, stability filtering, and QSAR-ready interpretation. The normalized occupancy statistic was

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)]} \qquad (298)$$

Here, c denotes an interaction class, S denotes receptor or structural states, T denotes time or sampling indices, and $I_{c,s,t}^{(k)}$ indicates whether interaction class c was observed for candidate k. Higher persistence indicates recurrent candidate-conditioned interaction events rather than one-pose artifacts.

This matters because one of the central risks in docking-driven screening is the overinterpretation of a single pose. The persistence statistic reduces this risk by asking whether interaction classes recur across states or sampling windows. SI Appendix I Figure 4 presents this persistence signal as a biology-facing stability/consistency cue that can support or challenge docking-driven promotions. This is consistent with the broader virtual-screening literature, where docking and scoring functions are repeatedly treated as useful but imperfect ranking devices rather than definitive affinity estimators [9–16].

The MBHA-guided update was

$$D\_B(\varrho_i,\varrho_j) = \surd(2 - 2\surd F(\varrho_i,\varrho_j)) \tag{299}$$

where x is the current candidate state, X_g is the best or elite aggregate, and $\odot$ denotes elementwise mixing. In the Results interpretation, this update is a controlled search heuristic, not a chemical-fusion rule. This distinction preserves the chemistry-safe interpretation of the HSX/HMC representation. It allows exploration around elite candidate neighborhoods while avoiding unsupported claims that heterogeneous pharmacophore modules are synthetically mergeable.

The biology-facing composite decision score was

$$F(\varrho_i,\varrho_j) = \mathrm{Tr}\surd(\surd\varrho_i\ \varrho_j\ \surd\varrho_i) \tag{300}$$

This expression reinforces the manuscript’s “fused signal, not a single metric” stance. Persistence, docking energy, regularization, QSAR interpretability, and action cost are separated rather than collapsed into an unexplained number. The same principle underlies the SI-derived refinement audit, where

$$g_{a\beta}(_x) = \langle\partial_a\varrho(x),\partial_\beta\varrho(x)\rangle HS \tag{301}$$

and

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{302}$$

The coefficients in this SI-derived audit are fixed for reproducible reporting. They should be changed only if an alternative weighting protocol is declared before interpretation. This is consistent with the manuscript’s audit-forward philosophy: model weights, rank shifts, diagnostic penalties, and topology supports should be logged and replayable.

The result is that the biology-facing and mathematical-chemistry panels support a reproducible prioritization hypothesis. They do not assert binding affinity, efficacy, toxicity, or clinical performance. Taken together, the Results show that GLYBATOMAQ™ converts a docking-derived candidate list into a structured rank-audit object. The strongest repeated signal was associated with the HSX-cons2/CONS2 family, which ranked first in the AutoDock Vina best-pocket analysis, first in the DockThor 3Q2R panel where visible, and first in the SI-derived DFT/QMC/MQWalk final fused score. However, because the AutoDock Vina separation between HSX-CONS2 and HSX-cons4 was only 0.1 kcal mol $^{-1}$, this ordering should be treated as rank-sensitive rather than definitive.

The DFT proxy branch supplied deterministic descriptor evidence for near-tie separation. The QMC-style branch supplied uncertainty-aware energetic auditing. MQWalk supplied topology-consistency validation on the PSD operator-overlap graph. Residue-level fingerprints showed that candidates with similar total-energy scale may still have different stabilizing and destabilizing contact patterns. The HSX/HMC representation preserved chemistry-safe family comparability without claiming atom-exact scaffold fusion. The persistence and MBHA/QSAR panels added stability and interpretability hooks without converting computational ranking into efficacy claims. Figure 10 provides the geometric rationale for this discussion by showing that the HSX/HMC candidates are best read as feasible pharmacophore paths constrained by role order, kernel compatibility, MQWalk support, and MBHA projection, rather than as unconstrained chemical fusions.

The final result is therefore a conservative prioritization statement:

$$\Pi_{curv} = \lambda_9 \| g - g_0 \| F^2 + \lambda\kappa \| \kappa - \kappa_0 \|_2^2 \quad (303)$$

This statement should be accompanied by the following boundary:

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \succeq 0,\ \mathrm{Tr}(\varrho_x) = 1 \quad (304)$$

Within that boundary, the method has a clear mathematical-chemistry contribution. It makes ligand rank movement explainable through

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \quad (305)$$

and constrains that movement through

$$\bar{K}(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)]} \quad (306)$$

Equations (267) and (268) summarize the research-level mathematical-chemistry usefulness of the Results section. The manuscript does not ask readers to trust a complex score. It shows how a rank changed, which evidence supported the change, which uncertainty constrained it, and whether the movement remained compatible with the operator-kernel topology shown across SI Appendix I Figure 1, SI Appendix I Figure 2, SI Appendix I Figure 3, SI Appendix I Figure 4, SI Appendix I Figure 5, SI Appendix I Figure 6, SI Appendix I Figure 7, SI Appendix I Figure 8, and SI Appendix I Figure 9.

3.11. Quantum-Geometry Reporting Outputs

The quantum-geometry reporting module converts the formal definitions of Section 2.1.10 into auditable output objects. In practical terms, a candidate is not promoted solely because its fused score improves. It must also have an interpretable PSD-overlap neighborhood, acceptable Bures-distance and projection residuals, bounded geodesic/curvature penalties, and sufficient MQWalk topology support. The expected reporting fields are defined in Tables 12 and 13.

Figures 11-14 provide the corresponding visual audit trail. Figure 11 identifies where DFT, QMC, MQWalk, and quantum-geometry diagnostics enter the workflow. Figure 12 clarifies that HMC/HSX candidates are paths on a constrained pharmacophore chart. Figure 13 gives a normalized matrix template for descriptor agreement, and Figure 14 shows how graph support is evaluated before rank movement is accepted. Together with Tables 14 and 15, these additions make the results section more transparent and reproducible as a quantum-geometric rank-audit framework.

No additional biological or clinical claim is introduced by these additions. The tables and figures strengthen reproducibility by specifying what must be reported for a candidate-level quantum-geometry audit: metric terms, distance terms, curvature terms, uncertainty terms, topology terms, projection residuals, and rank-shift consequences.

3.8. SI-calibrated QMC-DFT-MQWalk numerical enrichment

To keep the numerical reporting traceable, the final results layer was recalculated from the SI-derived branch summaries and the reproducibility logic in SI Appendix II. The quantities below are normalized proxy scores, not experimental thermodynamic values. They summarize how docking-oracle support, DFT proxy descriptor support, QMC reliability, MQWalk propagated support, Bures control, and diagnostic penalties combine under a classical quantum-geometric display formalism. The corresponding values expand Tables 10 and 12-15 and provide explicit numerical anchors for Figures 11-14.

$$N^E(L) = [E_{max} - E(L)] / [E_{max} - E_{min} + \varepsilon] \quad (307)$$

$$N^DFT(L) = [S^DFT(L) - \min S^DFT] / [\max S^DFT - \min S^DFT + \varepsilon] \quad (308)$$

$$NQMC(L) = rQMC(L)\ [1 - CIQMC(L)] \quad (309)$$

$$NMQW(L) = [SMQW(L) - \min SMQW] / [\max SMQW - \min SMQW + \varepsilon] \quad (310)$$

$$CB(L) = 1 - DB(L) \quad (311)$$

$$Snum(L) = 0.20\ N^E + 0.25\ N^DFT + 0.20\ NQMC + 0.25\ NMQW - \Pi diag \quad (312)$$

$$\mathrm{SQG,num}(L) = 0.22\ \mathrm{OPSD} + 0.16\ \mathrm{CB} + 0.24\ \mathrm{RDFT/QMC} + 0.24\ \mathrm{SMQW} + 0.14\ \mathrm{Fready} \tag{313}$$

$$\Delta S^{D}\mathrm{FT}(L) = 0.25\ N^{D}\mathrm{FT}(L) \tag{314}$$

$$\Delta \mathrm{SQMC}(L) = 0.20\ \mathrm{NQMC}(L) \tag{315}$$

$$\Delta \mathrm{SMQW}(L) = 0.25\ \mathrm{NMQW}(L) \tag{316}$$

$$\Pi\mathrm{diag}(L) = \lambda D[1 - \mathrm{rDFT}(L)] + \lambda Q[1 - \mathrm{rQMC}(L)] + \lambda T\ \Delta\mathrm{topo}(L) \tag{317}$$

$$\Delta\mathrm{topo}(L) = 1 - \mathrm{SMQW}(L) \tag{318}$$

$$\mathrm{Scert}(L) = \mathrm{SQG,num}(L) - 0.10\ \mathrm{DB}(L) \tag{319}$$

$$\mathrm{Rnum}(L) = \mathrm{rankL}[-\mathrm{Snum}(L)] \tag{320}$$

$$\mathrm{RQG}(L) = \mathrm{rankL}[-\mathrm{SQG,num}(L)] \tag{321}$$

$$\Delta\mathrm{Rnum}(L) = \mathrm{R0}(L) - \mathrm{Rnum}(L) \tag{322}$$

$$\mathrm{Meff,rel}(L) = 1 / [1 + \sigma\mathrm{QMC}^{2}(L)] \tag{323}$$

$$\mathrm{CIrel}(L) = 1 - \mathrm{CIQMC}(L) \tag{324}$$

$$\mathrm{QMCready}(L) = 1\{\mathrm{NQMC}(L) \geq 0.80\} \tag{325}$$

$$\mathrm{DFTready}(L) = 1\{N^{D}\mathrm{FT}(L) \geq 0.50\} \tag{326}$$

$$\mathrm{MQWready}(L) = 1\{\mathrm{SMQW}(L) \geq 0.75\} \tag{327}$$

$$\mathrm{Buresready}(L) = 1\{\mathrm{DB}(L) \leq 0.50\} \tag{328}$$

$$\mathrm{Fusionready}(L) = 1\{\mathrm{SQG,num}(L) \geq 0.65\} \tag{329}$$

$$\mathrm{Auditready}(L) = \mathrm{DFTready} \cdot \mathrm{QMCready} \cdot \mathrm{MQWready} \cdot \mathrm{Buresready} \cdot \mathrm{Fusionready} \tag{330}$$

$$\mathrm{Proxyoverlay}(L) = \mathrm{mean}[N^{E}, N^{D}\mathrm{FT}, \mathrm{NQMC}, \mathrm{NMQW}, \mathrm{CB}] \tag{331}$$

$$\text{Target-count ratio} = 48 / 234 = 0.205 \tag{332}$$

$$\text{GBM-overlap ratio} = 48 / 892 = 0.054 \tag{333}$$

$$\Delta G(\text{AP-6}) - \Delta G(\text{ilomastat}) = -13.7 - (-13.1) = -0.6\ \mathrm{kcal\ mol^{-1}} \tag{334}$$

$$\Delta G(\text{AP-7}) - \Delta G(\text{ilomastat}) = -13.3 - (-13.1) = -0.2\ \mathrm{kcal\ mol^{-1}} \tag{335}$$

$$\Delta\mathrm{Vina}(\mathrm{AKT1,HIF1A}) = -10.1 - (-5.6) = -4.5\ \mathrm{kcal\ mol^{-1}} \tag{336}$$

$$b(L) = [N^{E}, N^{D}\mathrm{FT}, \mathrm{NQMC}, \mathrm{NMQW}, \mathrm{CB}, \mathrm{SQG,num}] \tag{337}$$

$$||b(L)||_{2} = [\Sigma k\ bk(L)^{2}]^{1/2} \tag{338}$$

$$\mathrm{Rrobust}(L) = \mathrm{median}[b(L)] - \mathrm{MAD}[b(L)] \tag{339}$$

$$\mathrm{MarginQG}(Li,Lj) = \mathrm{SQG,num}(Li) - \mathrm{SQG,num}(Lj) \tag{340}$$

$$\mathrm{Promote}(L) \Leftrightarrow \Delta\mathrm{Rnum}(L) > 0 \text{ and } \mathrm{Auditready}(L) = 1 \tag{341}$$

$$\mathrm{Flag}(L) \Leftrightarrow \mathrm{Auditready}(L) = 0 \text{ or } \Pi\mathrm{diag}(L) > 0.05 \tag{342}$$

$$\mathrm{DivQG}(L) = |\mathrm{Snum}(L) - \mathrm{SQG,num}(L)| \tag{343}$$

$$\mathrm{Sconsensus}(L) = 0.50\ \mathrm{Snum}(L) + 0.50\ \mathrm{SQG,num}(L) \tag{344}$$

$$\mathrm{Rconsensus}(L) = \mathrm{rankL}[-\mathrm{Sconsensus}(L)] \tag{345}$$

$$\text{Numerical audit output} = \{\text{Tables 10, 16, 17, 18; Figures 11-14; SI I; SI II}\} \tag{346}$$

Table 16. SI-derived DFT/QMC/MQWalk numerical branch enrichment and recalculated fusion score.

Candidate	Base support	DFT proxy	QMC reliability	MQWalk support	Penalty	S_num (Eq. 496)	Rank
HSX-cons2	1.000	1.000	0.980	0.922	0.064	0.812	1
HSX-cons1	0.462	0.589	0.934	0.833	0.029	0.606	2
HSX-cons3	0.108	0.305	0.924	0.778	0.032	0.445	3
HSX-cons4	0.282	0.013	0.882	0.780	0.034	0.397	4

Note: Values are normalized branch proxies derived from the manuscript branch table and recalculated using Eq. (312). The terms follow the Algorithm 1 fusion logic in SI Appendix II and are interpreted as rank-audit support rather than experimental affinity.

Table 17. Quantum-geometric numerical reporting matrix with Bures-control and fusion-readiness proxies.

Candidate	PSD overlap	D_B distance	Bures control	DFT/QMC reliability	MQWalk support	Fusion readiness	S_QG,num (Eq. 497)	QG rank
HSX-cons1	0.88	0.30	0.70	0.72	0.86	0.78	0.794	2
HSX-cons2	0.95	0.24	0.76	0.83	0.91	0.84	0.866	1
HSX-cons3	0.71	0.45	0.55	0.64	0.78	0.67	0.679	3
HSX-cons4	0.68	0.50	0.50	0.61	0.74	0.63	0.642	4

Note: Values are calculated from the SI-aligned reporting matrix used in Figures 12 and 13. DB is reported as a normalized distance proxy, while Bures control is calculated as 1 - DB before fusion. These values enrich the QMC-DFT-MQWalk audit with comparable numerical quantum-geometry channels.

Table 18. SI Appendix I and II numerical anchors used to calculate the enriched quantum-geometry proxy values.

SI numerical anchor	Value	Calculated diagnostic	Result	Used in
Vortioxetine unique targets	234	Shared/unique target ratio	48/234 = 0.205	Figure 11; Discussion
Glioblastoma disease proteins	892	Shared/GBM protein ratio	48/892 = 0.054	Figure 11; Results
STRING PPI network	47 nodes / 255 edges	Refined PPI reduction	17 nodes / 85 edges	Figure 11; Table 18
Core PPI network	5 targets / 10 edges	Core-to-shared target ratio	5/48 = 0.104	Discussion
AKT1 Vina score	-10.1 kcal mol^{-1}	Span vs HIF1A	-4.5 kcal mol^{-1}	Figure 11
MMP3 AP-6 MD ΔbindG	-13.7 ± 0.4	Gap vs ilomastat	-0.6 kcal mol^{-1}	Figure 12
MMP3 AP-7 MD ΔbindG	-13.3 ± 0.4	Gap vs ilomastat	-0.2 kcal mol^{-1}	Figure 12
SI Appendix II fusion inputs	w_0,w_D,w_Q,w_M, penalties	Operational output	Tables 16-17	Methods/Results

Note: The table records only SI-derived or SI-calculated numerical anchors. They are used to contextualize rank-audit behavior and do not constitute experimental binding, efficacy, or clinical validation.

Tables 16-18 extend the numerical audit beyond the visual reporting matrix. Table 16 recalculates DFT, QMC, MQWalk, and penalty-weighted support from the branch-score audit, Table 17 converts the PSD/Bures/DFT-QMC/MQWalk channels into an explicit quantum-geometric fusion score, and Table 18 records the SI Appendix I and II anchors used to calculate the numerical context. These tables are cited in the Discussion to distinguish reproducible rank-audit values from experimental binding claims.

4. Discussion

The enriched numerical layer in Tables 16-18 strengthens the quantum-geometric interpretation by placing DFT proxy support, QMC reliability, MQWalk topology support, Bures control, fusion readiness, and SI-derived biological anchors into a common normalized scale. These values support the figure-based interpretation in Figures 11-14 and provide a quantitative bridge between the SI Appendix I numerical panels and the SI Appendix II reproducibility algorithms.

The Discussion interprets the GLYBATOMAQ™ framework as an original quantum-geometric rank-audit formalism. The central mathematical objects are the normalized PSD descriptor operator, the Bures-style control distance, the Fubini-Study/pullback metric, the quantum Fisher sensitivity, Berry-type gauge consistency, MQWalk support, and reliability-weighted fusion, as summarized in Table 12 and visualized in Figures 10-14. The discussion therefore follows the structure of the manuscript figures and tables as a continuous original-research argument.

4.1. Interpretation of GLYBATOMAQ™ as a Rank-Audit Framework, not a Docking Engine

GLYBATOMAQ™ should be interpreted as a **post-docking, pattern-recognition, and rank-audit framework**, not as a docking engine and not as a physical simulator of protein–ligand binding. This distinction is central to the manuscript. The purpose of GLYBATOMAQ™ is not to claim that a computational score equals binding affinity, biological efficacy, or therapeutic performance. Its purpose is to ask a narrower and more mathematically controlled question: **when a ligand changes rank after refinement, can that rank movement be decomposed into traceable electronic, uncertainty, and topology-consistency evidence?**

The structural setting is anchored to the Protein Data Bank [1], the human GLIPR1 structural study of Asojo et al. [2], and the GLIPR1-related structures 3Q2U [3] and 3Q2R [4], with pocket definition supported by Fpocket [5]. These structures define the receptor-side context in which ligand prioritization is performed. However, GLYBATOMAQ™ does not treat a selected structure or pocket as a complete biological model. Instead, each pocket defines a protocol-specific ranking environment. The output is therefore pocket-aware and rank-aware, not an absolute statement about molecular recognition.

Upstream docking is treated as a **fixed-protocol oracle**. Engines such as AutoDock Vina [6], AutoDock4 [7], Glide [8], and DockThor-VS [13] are used to generate comparative signals, not literal experimental affinities. This interpretation is consistent with the docking literature. Kitchen et al. noted that "significant challenges remain" in docking and scoring [9], Warren et al. emphasized careful benchmark-based interpretation of docking programs and scoring functions [10], and later reviews reinforced that docking is most defensible as a comparative prioritization tool rather than as a direct affinity estimator [11,12]. Benchmark resources such as PDBbind [14], the docking sets of Huang et al. [15], and DUD-E [16] likewise emphasize the rank-sensitive nature of virtual screening.

Within this interpretation, the first contribution of GLYBATOMAQ™ is conceptual discipline. Docking defines the upstream frontier,

$$\text{D\_B}(\varrho_i,\varrho_j) = \sqrt{}(2 - 2\sqrt{}\text{F}(\varrho_i,\varrho_j)) \tag{347}$$

but the final claim is not that the top docking score is biologically true. The final claim is that subsequent rank movement,

$$\text{F}(\varrho_i,\varrho_j) = \text{Tr}\sqrt{}(\sqrt{}\varrho_i\ \varrho_j\ \sqrt{}\varrho_i) \tag{348}$$

can be audited. This rank-first logic is summarized in SI Appendix I Figure 1 and SI Appendix I Figure 2, where docking/oracle scoring is stabilized by an operator-kernel layer before DFT, QMC, MQWalk validation, and reliability-aware fusion.

This framing formally resolves the possibility that the workflow might appear uncontrolledly complicated. The method is not intended to impress through multiple advanced labels. It is intended to prevent a common failure in computational screening: the conversion of an unstable docking table into an overconfident biological story. GLYBATOMAQ™ instead produces a structured audit trail.

4.2. Mathematical-Chemistry Basis: Structured Ranking Rather than Scalar Rescoring

The quantum-geometry figures and tables sharpen this mathematical-chemistry basis. Figure 11 and Table 12 define the operator-manifold objects that convert ligand prioritization from scalar rescoring into geometry-constrained rank auditing. Figure 13 and Table 13 then translate those objects into a candidate-level reporting matrix, making the quantum-geometric assumptions explicit rather than implicit.

The primary mathematical-chemistry contribution of GLYBATOMAQ™ is the recasting of ligand prioritization as a **structured ranking problem** rather than as a scalar rescoring exercise. Conventional docking workflows often produce a list of candidates sorted by a score. GLYBATOMAQ™ preserves that list as the initial frontier but then asks whether rank changes are supported by additional mathematical evidence.

The core workflow can be written as

$$g_{a\beta}(_x) = \langle \partial_a\varrho(x),\partial_\beta\varrho(x)\rangle \text{HS} \tag{349}$$

The object of interest is not merely

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{350}$$

but the full audit bundle

$$\Pi_{\mathrm{curv}} = \lambda_9 \| g - g_0 \| F^2 + \lambda\kappa \| \kappa - \kappa_0 \|_2^2 \tag{351}$$

This formulation makes the method useful to mathematical chemistry because it turns ligand prioritization into a reproducible transformation with interpretable intermediate objects. A ligand is not simply "better" because it has a favorable score. It is promoted only if the promotion survives electronic, uncertainty, and topology checks.

This is particularly important in near-tie regions. If two candidates differ only slightly in docking score, then a raw leaderboard may imply a stronger ordering than the calculation justifies. GLYBATOMAQ™ treats such near ties as cases requiring additional evidence. The DFT branch asks whether deterministic electronic descriptors support the separation. The QMC branch asks whether energetic support survives uncertainty. The MQWalk branch asks whether the candidate remains coherent in the operator-overlap graph. The final rank shift is then reported with reliability terms and flags.

Thus, the framework is not merely a score generator. It is a **mathematical decision audit**.

4.3. Chemistry-Safe Molecular Representation and the HMC Consensus Abstraction

The molecular representation layer remains deliberately chemistry-safe. It is compatible with MMFF94 geometry handling [17], RDKit [18], Morgan-style encoding [19], ECFP fingerprints [20], SMILES [21], and SELFIES [22]. It is also conceptually compatible with grammar VAE [23], junction-tree VAE [24], graph-policy learning [25], recurrent molecular design [28], latent molecular design [29], neural message passing [30], graph convolution [31], and attention-based representation learning [32].

The role of these representations is not to claim that a generated or consensus object is necessarily synthesizable, active, or clinically relevant. Their role is to create controlled molecular encodings that support comparison, clustering, shortlisting, and audit. This distinction is especially important for the Hyper Mega Core concept. The HMC is not an atom-exact mega-scaffold and should not be read as a forced chemical fusion across incompatible chemotypes. Instead, it is a **numerical consensus topology**.

In Figure 10, the HMC is represented as a chemistry-safe quantum-geometric consensus geometry for docking-oriented pattern recognition and audit-forward re-ranking. Nodes encode pharmacophore-role modules such as recognition nucleus, linker, interaction node, spacer, polarity/solubility module, and optional tail/cap. Edges represent descriptor-space compatibility, while the enriched construction assigns each module a normalized PSD operator and tests admissibility through trace overlap, Bures/Fubini-Study-style distance, curvature consistency, and MQWalk propagation support. The figure explicitly frames these links as ranking, clustering, and topology-audit aids, not synthetic fusion rules.

Figures 11 and 12 extend this interpretation from the HMC manifold to the full quantum-geometric audit pipeline. Figure 11 emphasizes that the bounded hypergeometric channel, DFT descriptor branch, QMC uncertainty branch, and MQWalk graph branch are fused only after normalization and reliability gating. Figure 12 generalizes the same logic to proxy overlays that include SI Appendix I biological and chemical anchors, such as glioblastoma target-harvesting counts, core target docking scores, MMP3 inhibitor ranking anchors, and HSX/comparator docking-energy proxies. The discussion therefore uses the main figures as compact interpretive maps and the SI Appendix I figures as transparent numerical provenance.

This representation is important because it separates **mathematical comparability** from **chemical literalism**. GLYBATOMAQ™ can compare families of candidates without claiming that those families collapse into one physical scaffold. That distinction strengthens the manuscript because it prevents overinterpretation of the HMC and clarifies its role as a pattern-recognition object.

4.4. Positive-Semidefinite Operator Kernels as the Comparability Layer

A central mathematical component of GLYBATOMAQ™ is the positive-semidefinite operator representation. Candidate comparison is formalized through operators

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \geq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{352}$$

typically constructed as

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{353}$$

This construction places pocket and ligand descriptors into a controlled operator space. The trace-overlap similarity is

$$\bar{K}(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{}[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)] \tag{354}$$

and the normalized overlap kernel is

$$\mathrm{D\_B}(\varrho_i,\varrho_j) = \sqrt{}(2 - 2\sqrt{}F(\varrho_i,\varrho_j)) \tag{355}$$

This normalized kernel has a clear role: it stabilizes ligand–pocket comparability before refinement. Without this layer, DFT, QMC, docking, hotspot, and residue-contact quantities would be difficult to compare coherently. The operator-kernel construction is consistent with kernel learning [33], support-vector methods [34], Gaussian-process reasoning [35], and positive-definite matrix analysis [36].

The normalized kernel also supplies the graph used for MQWalk topology validation. Thus, the operator layer is not an optional non-functional formalism. It connects representation, ranking, topology, and validation. SI Appendix I Figure 2 shows this kernel layer as the central PSD/operator stage, and SI Appendix I Figure 6 shows how the same normalized overlap graph supports MQWalk/teleport validation.

This is one of the most important mathematical-chemistry contributions of the workflow. It replaces a purely scalar docking table with an operator-geometric comparison space.

Table 12 and Figure 13 define the manuscript-wide naming convention for these quantum-geometric objects. They clarify that normalized PSD operators, Bures-style distance, pullback geometry, quantum Fisher sensitivity, Berry-type gauge consistency, curvature penalties, and MQWalk support have distinct computational meanings. This prevents the Results and Discussion sections from collapsing all quantum-geometric objects into a single scalar score.

4.5. Avogadro-Conditioned Stabilization as Numerical Regularization, not Physical Overclaiming

Another component requiring careful interpretation is the Avogadro-conditioned stabilization map. GLYBATOMAQ™ uses robust feature conditioning,

$$F(\varrho_i,\varrho_j) = \mathrm{Tr}\sqrt{}(\sqrt{}\varrho_i\ \varrho_j\ \sqrt{}\varrho_i) \tag{356}$$

followed by

$$g_{a\beta}(x) = \langle \partial_a \varrho(x), \partial_\beta \varrho(x) \rangle \mathrm{HS} \tag{357}$$

and

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{358}$$

Here N_A is the CODATA Avogadro constant [43]. Its role is numerical, not physical. The manuscript does not claim that this scaling represents molecular-counting physics in the screening process. Instead, N_A is used as a deterministic scale anchor that supports reproducible damping across heterogeneous branches.

This distinction is essential. Without robust normalization, the fusion rule could be dominated by one large-magnitude descriptor or by an unstable outlier. The transformation

$$\Pi_{\mathrm{curv}} = \lambda_9 \| g - g_0 \| F^2 + \lambda\kappa \| \kappa - \kappa_0 \|_2^2 \tag{359}$$

ensures that branch features are centered, damped, and bounded before fusion. This is compatible with the manuscript's approximation-theoretic basis [37-42] and with the practical need to make docking, DFT, QMC, topology, and residue-contact descriptors comparable.

SI Appendix I Figure 2 and SI Appendix I Figure 5 show this conditioning stage as part of the rank-centric refinement circuit. Its purpose is regularization and reproducibility, not physical embellishment.

4.6. Complementarity of DFT and QMC Refinement

The DFT and QMC branches are complementary rather than redundant.

The DFT branch provides deterministic electronic-structure descriptors grounded in the Hohenberg–Kohn theorem [44], the Kohn–Sham equations [45], and standard DFT treatments [46,47]. For a logged geometry $\chi^{\star}(P, L)$, the branch derives a descriptor vector

$$\varrho_x = X_x X_x^{T} / \mathrm{Tr}(X_x X_x^{T}),\ \varrho_x \succeq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{360}$$

and maps it to

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{361}$$

The DFT branch is not used as an absolute binding-energy calculator. It functions as a deterministic electronic descriptor generator for near-tie discrimination. Its contribution is reliability-gated through

$$\bar{K}(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)]} \tag{362}$$

Therefore, SCF failures, unstable geometries, incomplete descriptors, or out-of-domain electronic features are not silently absorbed into the final score. They reduce reliability or trigger diagnostic penalties.

The QMC branch has a different role. It provides selective stochastic energetic auditing, consistent with QMC methodology [48,49]. Its output is summarized by

$$D\_B(\varrho_i,\varrho_j) = \sqrt{(2 - 2\sqrt{F}(\varrho_i,\varrho_j))} \tag{363}$$

and its conservative score is

$$F(\varrho_i,\varrho_j) = \mathrm{Tr}\sqrt{(\sqrt{\varrho_i}\ \varrho_j\ \sqrt{\varrho_i})} \tag{364}$$

This structure ensures that a favorable energetic estimate does not automatically produce a rank promotion if uncertainty is large. The QMC branch therefore answers a question that DFT does not: **is the apparent energetic support stable under stochastic uncertainty?**

The two branches can be summarized as follows:

$$g_{a\beta}(_x) = \langle \partial_a \varrho(x), \partial_\beta \varrho(x) \rangle HS \tag{365}$$

Together, they make near-tie refinement more defensible. SI Appendix I Figure 5 visualizes the DFT descriptor distributions and the fusion circuit, while SI Appendix I Figure 2 places DFT and QMC as distinct late-stage refinement branches.

4.7. MQWalk as Topology Validation, not Quantum-Biological Transport

Figure 14 makes the MQWalk interpretation visually explicit: propagated mass is a graph-topology diagnostic over normalized operator-overlap weights. It is not evidence for coherent biomolecular quantum transport. The purpose of the MQWalk branch is to reject or penalize rank promotions that are favorable electronically but unsupported by the operator-neighborhood geometry.

MQWalk is perhaps the component most likely to be misunderstood. It uses notation inspired by continuous-time quantum walks and quantum-information formalisms [50], but it is not presented as a physical model of quantum transport in GLIPR1 or ligand recognition. This caution is necessary given broader limits of overextending quantum language to biology [51,52].

In GLYBATOMAQ™, MQWalk operates on the normalized operator-overlap graph. From $\bar{K}$, a bipartite adjacency is constructed:

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{366}$$

The degree-normalized adjacency is

$$\Pi_{\mathrm{curv}} = \lambda_9 \| g - g_0 \| F^2 + \lambda\kappa \| \kappa - \kappa_0 \|_2^2 \tag{367}$$

doi:10.20944/preprints202605.2096.v1

Teleport-style stabilization gives

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \succeq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{368}$$

The Avogadro-conditioned Laplacian is

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{369}$$

and the propagation operator is

$$\tilde{K}(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{}[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)] \tag{370}$$

Node mass is written as

$$D\_B(\varrho_i,\varrho_j) = \sqrt{}(2 - 2\sqrt{}F(\varrho_i,\varrho_j)) \tag{371}$$

These quantities are interpreted only as graph-propagation diagnostics. The MQWalk score is

$$F(\varrho_i,\varrho_j) = \mathrm{Tr}\sqrt{}(\sqrt{}\varrho_i\ \varrho_j \sqrt{}\varrho_i) \tag{372}$$

The term $\Delta_{\text{topo}}(P,L)$ penalizes or flags topology-discordant promotions. This means that a candidate promoted by DFT or QMC can still be questioned if it is unsupported by the operator-overlap neighborhood.

This role is shown in SI Appendix I Figure 1, SI Appendix I Figure 3, and SI Appendix I Figure 6, where MQWalk-style transport traces, heatmaps, and topology diagnostics are presented as consistency checks rather than physical observables.

Thus, MQWalk is not a claim about quantum biology. It is a topology-validation filter. Its value lies in constraining unsupported rank jumps.

4.8. Reliability-Aware Fusion, Diagnostic Penalties, and Rank-Audit Interpretation

The final GLYBATOMAQ™ score is retained as a reliability-aware evidence-integration rule rather than as a black-box ensemble. Its purpose is to preserve the upstream docking and pattern-recognition frontier while allowing only reliable electronic, stochastic-energetic, and topology-consistency evidence to influence post-refinement ranking. For a protein pocket P and ligand L, the fused score is defined as

$$g_{a\beta}(x) = \langle \partial_a \varrho(x), \partial_\beta \varrho(x) \rangle HS \tag{373}$$

The branch weights w_0, w_{DFT}, w_{QMC}, and w_{MQW} are predeclared or recorded in the protocol file. The quantities r_{DFT} and r_{QMC} are reliability gates with values in $[0, 1]$. These gates prevent unstable electronic-structure calculations or poorly converged stochastic estimates from being assigned the same evidentiary status as stable calculations. The total diagnostic penalty is additive:

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{374}$$

This decomposition is central to the audit-forward interpretation of the method. A candidate is not promoted simply because one scalar subscore is favorable. Promotion is supported only when the favorable signal is accompanied by adequate convergence, controlled uncertainty, compatible topology, and complete provenance.

The upstream contribution,

$$\Pi_{\text{curv}} = \lambda_9 \| g - g_0 \| F^2 + \lambda\kappa \| \kappa - \kappa_0 \|_2^2 \tag{375}$$

preserves the docking and pattern-recognition frontier. Thus, the refinement step remains a local audit of rank movement rather than an unconstrained replacement of the original screening order. The DFT contribution,

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \succeq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{376}$$

adds deterministic electronic-structure evidence only when the calculation is self-consistent, geometry-stable, descriptor-complete, and inside the calibration envelope of the shortlist. The QMC contribution,

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{377}$$

adds stochastic energetic support only after uncertainty has been penalized and sampling adequacy has been checked. The MQWalk contribution,

$$\tilde{K}(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) / \sqrt{}[\mathrm{Tr}(\varrho P^2)\ \mathrm{Tr}(\varrho L^2)] \tag{378}$$

adds topology-consistency support derived from the normalized positive-semidefinite operator-overlap graph. Finally, the penalty term,

$$D\_B(\varrho_i,\varrho_j) = \sqrt{}(2 - 2\sqrt{}F(\varrho_i,\varrho_j)) \tag{379}$$

ensures that instability, uncertainty inflation, topology discordance, and missing provenance remain visible in the final score rather than being hidden inside an apparently precise ranking. Therefore, $S_{\text{final}}^{(\mu m \backslash \text{ⓜ}_{+} \quad)}$ should be interpreted as a mathematically structured audit statement, not as a direct claim of experimental binding affinity.

For clarity, the post-refinement rank is obtained by sorting the fused score in descending order:

$$F(\varrho_i,\varrho_j) = \mathrm{Tr}\sqrt{}(\sqrt{}\varrho_i\ \varrho_j\ \sqrt{}\varrho_i) \tag{380}$$

where the sign convention may be reversed if the implementation stores lower-is-better energies. The pre-refinement rank is

$$g_{a\beta}(_x) = \langle \partial_a\varrho(x),\partial_\beta\varrho(x)\rangle HS \tag{381}$$

The rank-shift diagnostic is then

$$\kappa(x) = \mathrm{Ric}[g(x)] \tag{382}$$

A positive ΔR indicates promotion after refinement, a negative ΔR indicates demotion, and $\Delta R = 0$indicates rank preservation. Shortlist stability is summarized by

$$\Pi_{\text{curv}} = \lambda_9 \| g\text{-}g_0 \| F^2 + \lambda\kappa \| \kappa\text{-}\kappa_0 \|_2{}^2 \tag{383}$$

This statistic does not require the ranking to reproduce the upstream order. Instead, it identifies whether major reordering has occurred and whether such reordering is explained by DFT evidence, QMC evidence, or MQWalk topology support. A low Overlap@k may be chemically meaningful, but only if the branch-level audit shows that the displacement is supported and is not a consequence of numerical instability.

4.9. Quantitative Validation Framework for Rank-Audit Reproducibility

The referee correctly noted that the previous version of the manuscript did not provide a sufficiently explicit quantitative validation framework. In this presentation, validation is separated into four levels: external benchmark validation, internal SI-derived baseline validation, rank-correlation validation, and uncertainty-aware audit validation. This separation prevents the method from being evaluated only against itself and avoids conflating internal docking figures with fully curated active/decoy benchmarks.

For a curated benchmark set

$$\varrho_x = X_x X_x^T / \mathrm{Tr}(X_x X_x^T),\ \varrho_x \succeq 0,\ \mathrm{Tr}(\varrho_x) = 1 \tag{384}$$

where $\mathcal{A}$ is the set of known actives and $\mathcal{Z}$ is a matched decoy set, the primary binary classification metric is the receiver-operating-characteristic area under the curve:

$$K(P,L) = \mathrm{Tr}(\varrho P\ \varrho L) \tag{385}$$

Because virtual screening is most useful in early recognition, enrichment factors are also reported:

$$\Xi_{311}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5 \Pi\_diag \tag{386}$$

where α is the screened fraction, N_α is the number of compounds in the top α fraction, and Hits_α is the number of known actives recovered in that fraction. Early enrichment is further assessed using the Boltzmann-enhanced discrimination of ROC score:

$$\Xi_{312}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5 \Pi\_diag \tag{387}$$

where r_i is the rank of active compound i, and C_α is the standard normalization constant.

Rank agreement with an independent baseline b, such as docking-only or DockThor-only ordering, is measured by Spearman correlation,

$$\Xi_{313}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5 \Pi\_diag \tag{388}$$

and Kendall correlation,

$$\Xi_{314}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{389}$$

This manuscript should explicitly state that external benchmark validation requires curated GLIPR1 actives and inactives where available. If GLIPR1-specific experimental sets are insufficient, the study should report a two-level validation strategy: first, a target-specific GLIPR1 comparator analysis using known ligands, reference drugs, or protein-family-relevant compounds; and second, a method-level docking benchmark analysis on standard ligand-decoy sets, such as DUD-E or PDBbind-derived targets. The first level evaluates chemical plausibility in the GLIPR1 setting. The second level evaluates whether the rank-audit machinery behaves sensibly on recognized benchmarking infrastructure. These two validation tasks answer different questions and should not be collapsed.

The SI-derived internal validation is reported conservatively. SI Appendix I Figure 5B supplies the comparative docking-energy overview for HSX consensus candidates and FDA/reference comparators. SI Appendix I Figures 7 and 8 provide AutoDock Vina best-pocket and pocket-resolved baseline behavior. SI Appendix I Figure 9 supplies a DockThor-style baseline. These SI-derived results are not presented as a substitute for active/decoy enrichment. Instead, they show that the proposed fused score can be compared with docking-only and docking-style baselines under a fixed computational protocol.

The baseline family is defined as

$$\Xi_{315}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{390}$$

Here, B_{Vina} is the AutoDock Vina best-pocket leaderboard, $B_{\texttt{DockThor}}$ is the DockThor affinity/component-energy readout, B_{internal} is the SI Figure 5B comparator panel, B_θ is the upstream GLYBATOMAQ pattern-recognition frontier, and B_{full} is the reliability-aware fused model. A validation table therefore reports here ROC-AUC, $\mathrm{EF}_{1\%}$, $\mathrm{EF}_{5\%}$, BEDROC, Spearman ρ_s, Kendall τ, $\texttt{Overlap@}k$, and ΔR summaries for B_{full} against each appropriate baseline.

4.10. Ablation Analysis: Controlled Isolation of Each Module

The referee also requested an ablation analysis. This manuscript defines a controlled branch-removal design that tests whether each GLYBATOMAQ™ component contributes nonredundant information. The full fused model is compared with a sequence of ablated models.

The docking/pattern-recognition baseline is

$$\Xi_{316}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{391}$$

The DFT-only refinement model is

$$\Xi_{317}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{392}$$

The QMC-only refinement model is

$$\Xi_{318}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{393}$$

The topology-only validation model is

$$\Xi_{319}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{394}$$

The electronic/stochastic model without topology validation is

$$\Xi_{320}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{395}$$

The reliability-disabled model is

$$\Xi_{321}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{396}$$

The penalty-disabled model is

$$\Xi_{322}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{397}$$

The full model and the Table 9 ablation matrix are anchored by Eq. (1). The expected diagnostic pattern is not that every branch improves every metric in every case. Rather, a useful branch should improve at least one predeclared validation dimension without causing unacceptable degradation in another. DFT is expected to improve near-tie discrimination and electronic interpretability. QMC is expected to reduce unsupported promotions under high uncertainty. MQWalk is expected to reduce topology-discordant promotions. Reliability gates are expected to prevent numerically unstable

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2026 doi:10.20944/preprints202605.2096.v1

candidates from receiving unjustified rank gains. Diagnostic penalties are expected to increase audit transparency and reduce false confidence.

For each ablation model M, rank shift relative to the upstream model is defined as

$$\Xi_{323}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5\Pi\_diag \quad (398)$$

The branch contribution to rank movement is measured by

$$\Xi_{324}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5\Pi\_diag \quad (399)$$

where denotes the full model with branch removed. If is large for a candidate, the topology branch materially affects that candidate's rank. If is large only when confidence intervals are wide, the QMC branch is operating as an uncertainty control rather than as a universal score amplifier. This formulation formally resolves the reader's request because each module is evaluated as a removable hypothesis rather than as a non-functional part of a complex stack.

Table 9. Recommended ablation analysis framework for controlled module isolation.

Model	Terms retained	Question answered	Primary metrics	Expected interpretation
M_0	S_θ only	What does the upstream rank frontier achieve alone?	ROC-AUC, EF, BEDROC, ρ_s	Docking/pattern-recognition baseline.
M_{DFT}	S_θ + reliable DFT	Does deterministic electronic structure resolve near ties?	ΔR, near-tie accuracy, descriptor stability	Useful if DFT changes crowded regions without destroying the frontier.
M_{QMC}	S_θ + reliable QMC	Does uncertainty-aware energy prevent fragile promotions?	CI width, σ_E, uncertainty-weighted ΔR	Useful if high-uncertainty candidates are not overpromoted.
M_{MQW}	S_θ + topology	Does graph topology identify unsupported rank jumps?	Topology discordance, $\mathrm{Overlap}@k$	Useful if isolated candidates are flagged or demoted.
$M_{\mathrm{DFT+QMC}}$	Electronic and stochastic branches only	Is topology required beyond quantum refinement?	EF, ΔR, discordance	Assesses whether MQWalk adds nonredundant graph support.
$M_{\text{no-rel}}$	All scores without gates	Do reliability gates matter?	False-promotion rate, instability count	Should expose overpromotion of unreliable candidates.
$M_{\text{no-pen}}$	All gates without diagnostic penalty	Do penalties add audit value?	Penalty flags, ΔR, report completeness	Should expose hidden instability when penalties are removed.
M_{full}	All terms and penalties	Does the full chain improve rank auditability?	All metrics	Accepted only if performance and interpretability improve together.

4.11. Practical Role of MQWalk and Avoidance of Quantum Overstatement

The previous version contained detailed MQWalk formalism but did not sufficiently isolate its practical role. These states explicitly that MQWalk is a topology validator, not a physical model of biomolecular quantum transport. It operates on the same normalized operator-overlap graph that stabilizes the upstream representation.

The normalized pocket-ligand overlap is

$$\Xi_{325}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5\Pi\_diag \quad (400)$$

This overlap defines a bipartite adjacency matrix,

$$\Xi_{326}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5\Pi\_diag \quad (401)$$

The degree-normalized adjacency is

$$\Xi_{327}(L) = \alpha_1\bar{K} + \alpha_2 SDFT + \alpha_3 SQMC + \alpha_4 SMQW - \alpha_5\Pi\_diag \quad (402)$$

Teleport mixing is applied to prevent disconnected shortlist components from generating unstable propagation:

doi:10.20944/preprints202605.2096.v1

$$\Xi_{328}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{403}$$

The corresponding graph Laplacian is

$$\Xi_{329}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{404}$$

The walk-style propagation operator is

$$\Xi_{330}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{405}$$

and the node mass is

$$\Xi_{331}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{406}$$

The MQWalk score is computed as a balance between topology-supported persistence and discordance penalty:

$$\Xi_{332}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{407}$$

A practical discordance term can be written as

$$\Xi_{333}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{408}$$

where z_{branch} is the standardized DFT/QMC-supported promotion signal and z_{top} is the standardized topology-supported promotion signal. If DFT/QMC strongly promotes a ligand but the ligand is isolated in the operator-overlap graph, D_{top} becomes large and the promotion is penalized. Conversely, if a ligand is promoted by branch evidence and remains embedded in a coherent topology neighborhood, D_{top} is small and the promotion is considered topology-consistent.

This definition clarifies the practical role of MQWalk in three ways. First, MQWalk corrects rank errors when a single branch score promotes a topologically isolated candidate. Second, it helps prevent false positives by requiring rank movement to agree with neighborhood structure. Third, it improves robustness by showing whether a promotion persists under graph propagation rather than appearing only in a single scalar descriptor. The SI topology panels should therefore be described as inspection surfaces for dispersion, concentration, neighborhood persistence, and discordance, not as evidence of physical quantum transport.

4.12. Chemical and Biological Interpretation of HSX/HMC Candidates

This research explicitly states that HMC and HSX consensus representations are chemistry-safe reference geometries for comparison, clustering, docking-oriented pattern recognition, and re-ranking. They are not atom-exact "mega-scaffolds," synthesis prescriptions, or validated therapeutic structures. Descriptor-space proximity does not imply synthetic mergeability.

The HMC is represented by pharmacophore-role modules:

$$\Xi_{334}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{409}$$

Edges in the HMC graph represent descriptor-space compatibility rather than chemical bonds:

$$\Xi_{335}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{410}$$

An HSX consensus candidate is therefore an enumeration anchor,

$$\Xi_{336}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{411}$$

where Ω_{enum} records the allowed enumeration and chemistry-filtering rules. The intended status of $L_{\mathrm{HSX}}$is computational: it is a reproducible comparator object used to test whether the rank-audit pipeline can stabilize and interpret screening outputs. It is not reported as a confirmed binder, lead compound, or synthetic route.

Biological interpretation should remain restricted to pocket-conditioned hypotheses. The GLIPR1 structural context is anchored to experimentally determined sGLIPR1 structures, including 3Q2U and 3Q2R, and pocket definitions are treated as local chemical environments. For candidate L in pocket P, residue-resolved interpretation is summarized by

$$\Xi_{337}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{412}$$

where V_{res} is the residue/contact vocabulary and e_r is the extracted energetic or contact contribution. The sparse fingerprint is

$$\Xi_{338}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{413}$$

with hotspot support

$$\Xi_{339}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{414}$$

This formulation explains how two candidates may have similar total scores but different energetic wiring diagrams. SI Appendix I Figure 6 provides the intended form of interpretation: residue-resolved energetic fingerprints distinguish stabilizing and destabilizing contributions and connect numerical rank with structural contact patterns. The correct claim is therefore not that candidate *L* binds GLIPR1 experimentally. The correct claim is that candidate *L* is a pocket-conditioned computational hypothesis whose rank movement is supported, weakly supported, or contradicted by residue-level, uncertainty-level, and topology-level diagnostics.

4.13. Internal Numerical Audits

The SI-derived audit should be presented as an internal demonstration of the rank-audit logic. It should not be described as a complete external active/decoy benchmark. The internal result is nevertheless useful because it shows how the score integrates multiple rank-bearing inputs and how branch diagnostics can change the order of closely ranked candidates.

Table 10 reports the following normalized fields: base support, DFT proxy support, QMC reliability/support, MQWalk support, diagnostic penalty, final fused score, pre-rank, post-rank, and rank shift. In this audit, HSX-cons2 remains first after fusion, HSX-cons1 remains second, HSX-cons3 moves from fourth to third, and HSX-cons4 moves from third to fourth. This is a useful internal demonstration because the fused model does not simply copy one docking table. Instead, it integrates the base frontier, deterministic descriptor-style evidence, uncertainty-aware support, topology support, and diagnostic penalties.

The rank-transition matrix associated with the Table 10 internal audit is

$$\Xi_{340}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{415}$$

The corresponding rank-shift vector is

$$\Xi_{341}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{416}$$

The interpretation is deliberately modest. These rank shifts indicate that the audit framework can distinguish stable leaders from lower-tier near-tie candidates whose ordering is sensitive to branch evidence. They do not establish experimental activity. The proper conclusion is that the method creates a replayable audit trail for rank movement, not that it proves binding efficacy.

Table 10. Internal DFT/QMC/MQWalk branch-score audit and rank movement for HSX-cons1–4.

Candidate	Base support	DFT-style support	QMC reliability	MQWalk support	Penalty	Fused score	Rank movement
GLYBATOMAQ-HSX-cons2	1.000	1.000	0.980	0.922	0.064	0.852	1 → 1
GLYBATOMAQ-HSX-cons1	0.462	0.589	0.934	0.833	0.029	0.503	2 → 2
GLYBATOMAQ-HSX-cons3	0.108	0.305	0.924	0.778	0.032	0.235	4 → 3
GLYBATOMAQ-HSX-cons4	0.282	0.013	0.882	0.780	0.034	0.224	3 → 4

4.14. Reporting Standards and Reproducibility Checklist

Tables 14 and 15 were added to make figure/table citation and reproducibility checks explicit for submission. They define where each quantum-geometry visual is cited in the manuscript and what minimum fields should be reported for any candidate-level quantum-geometric GLYBATOMAQ™ audit.

To make it reproducible, every reported candidate should carry a rank-audit record:

$$\Xi_{342}(\mathrm{L}) = \alpha_1\bar{\mathrm{K}} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{417}$$

The provenance field must include the receptor structure, pocket definition, ligand identity, conformer source, protonation/tautomeric state, docking program and version, grid center and box size, scoring settings, DFT functional/basis or descriptor protocol, QMC sampling protocol where applicable, kernel construction, MQWalk graph threshold, teleport parameter, fusion weights, and any missing-value imputation flag. If this information is incomplete, Π_{prov} is nonzero.

Together with the notation glossary in Table 11, the manuscript includes the following reproducibility checklist: the compound library and benchmark set are identified; all baselines are named; all metrics are reported; all ablation models are run using identical splits and fixed weights; missing values are flagged; figures are interpreted as audit diagnostics rather than proof of binding; HMC/HSX objects are described as comparison anchors rather than synthetic claims; and experimental validation is stated to remain outside the computational claim.

4.15. Interpretation of Numerical Findings in Light of the Benchmark, Ablation, and Rank-Audit Framework

The numerical findings support the rank-centric and audit-forward interpretation of GLYBATOMAQ™ developed in the preceding sections. The purpose of the present evaluation is not to convert docking scores, DockThor affinity-like outputs, DFT proxy descriptors, QMC-style support terms, or MQWalk topology diagnostics into experimental binding affinities. Rather, the objective is to determine whether the proposed framework provides interpretable and non-redundant ranking evidence beyond a docking-only table. Accordingly, all numerical values in this section are interpreted as protocol-dependent computational signals for candidate prioritization, near-tie resolution, branch-level auditability, and topology-consistency analysis.

The benchmark and ablation structure can be expressed as a set of baseline models,

$$\Xi_{343}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{418}$$

where

$$\Xi_{344}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{419}$$

and

$$\Xi_{345}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{420}$$

This baseline decomposition is important because it prevents the full GLYBATOMAQ™ score from being evaluated only against itself. Instead, the final fused model is interpreted relative to docking-only baselines, an independent docking-style baseline, an internal comparator-ranking panel, and branch-ablated versions of the proposed framework.

For each ligand L_i and pocket P_j, the AutoDock Vina baseline can be written as

$$\Xi_{346}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{421}$$

with the best-pocket value defined by

$$\Xi_{347}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{422}$$

Under this baseline, GLYBATOMAQ-HSX-CONS2 was placed first with

$$\Xi_{348}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{423}$$

followed very closely by GLYBATOMAQ-HSX-cons4 with

$$\Xi_{349}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{424}$$

The absolute separation between these two leading Vina scores is therefore

$$\Xi_{350}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{425}$$

Because the separation in equation (390) is small relative to the practical uncertainty normally associated with docking-score interpretation, it should not be treated as a decisive biochemical distinction. Instead, it identifies a near-tie region of the ranking frontier. In the GLYBATOMAQ™ framework, such near ties are precisely the cases in which branch-level audit is most useful. Thus, equations (388)-(390) support the use of DFT/QMC/MQWalk refinement not as a replacement for docking, but as a controlled mechanism for interpreting rank-sensitive candidates.

The Vina-derived ordering can be represented as

$$\Xi_{351}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{426}$$

where lower docking energy corresponds to better rank. For the leading candidates,

$$\Xi_{352}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{427}$$

However, the near-tie condition can be expressed as

$$\Xi_{353}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{428}$$

where denotes a practical docking-score tolerance. When equation (393) holds, the ranking should be interpreted as uncertain or rank-sensitive rather than definitive. Therefore,

$$\Xi_{354}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{429}$$

where denotes an experimental dissociation constant. Equation (434) formalizes the conservative interpretation used throughout this manuscript: a favorable docking rank is a prioritization signal, not an experimental affinity measurement.

The comparative docking-energy overview provides a second internal baseline, B_2. In this analysis, the four HSX consensus candidates ranked above the listed FDA/reference comparators under the reported protocol. The ordering can be summarized as

$$\Xi_{355}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{430}$$

where “>” denotes a more favorable position in the reported docking-energy ranking. The corresponding reported values were

$$\Xi_{356}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{431}$$

and

$$\Xi_{357}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{432}$$

The strongest listed FDA/reference comparator in this panel was Vemurafenib, with

$$\Xi_{358}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{433}$$

Therefore, the separation between the weakest HSX candidate in this internal panel and the strongest listed FDA/reference comparator is

$$\Xi_{359}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{434}$$

Similarly, the separation between the strongest HSX candidate and the strongest listed FDA/reference comparator is

$$\Xi_{360}(L) = \alpha_1\bar{K} + \alpha_2\mathrm{SDFT} + \alpha_3\mathrm{SQMC} + \alpha_4\mathrm{SMQW} - \alpha_5\Pi\_\mathrm{diag} \tag{435}$$

Equations (435)-(442) indicate that the HSX consensus set occupies the favorable tier of the internal docking-energy overview. However, this panel is not a standard active/decoy benchmark. It does not by itself establish enrichment, prospective hit recovery, biological activity, or clinical relevance. Its appropriate interpretation is therefore

$$\Lambda_{361} = \mathrm{AggP}[\Xi_{361}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{436}$$

but

$$\Lambda_{362} = \mathrm{AggP}[\Xi_{362}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{437}$$

The DockThor 3Q2R panel provides an additional docking-style baseline, B_1. In that panel, GLYBATOMAQ-HSX-cons2 ranked first with

$$\Lambda_{363} = \mathrm{AggP}[\Xi_{363}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{438}$$

followed by Vorasidenib with

$$\Lambda_{364} = \mathrm{AggP}[\Xi_{364}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{439}$$

GLYBATOMAQ-HSX-cons1 with

$$\Lambda_{365} = \mathrm{AggP}[\Xi_{365}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{440}$$

and GLYBATOMAQ-HSX-cons4 with

$$\Lambda_{366} = \mathrm{AggP}[\Xi_{366}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{441}$$

The separation between Vorasidenib and HSX-cons1 in this DockThor panel is

$$\Lambda_{367} = \mathrm{AggP}[\Xi_{367}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{442}$$

and the separation between HSX-cons1 and HSX-cons4 is

$$\Lambda_{368} = \mathrm{AggP}[\Xi_{368}(P,L)] - \lambda\mathrm{BD\_B} - \lambda T\Delta_{\mathrm{topo}} - \lambda\mathrm{UCI\_QMC} \tag{443}$$

Equations (449) and (335) show that the second-to-fourth DockThor ranks form a close cluster. Thus, the DockThor baseline supports HSX-cons2 as a leading internal candidate while also confirming that several following candidates occupy another rank-sensitive region. This again supports the use of branch-level diagnostics rather than a single-score interpretation.

To connect the docking baselines to the GLYBATOMAQ™ refinement model, the pre-refinement rank is defined as

$$\Lambda_{369} = \mathrm{AggP}[\Xi_{369}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{444}$$

where S_{base} denotes the upstream docking/pattern-recognition support. The post-refinement rank is

$$\Lambda_{370} = \mathrm{AggP}[\Xi_{370}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{445}$$

The rank-shift diagnostic is then

$$\Lambda_{371} = \mathrm{AggP}[\Xi_{371}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{446}$$

A positive value of Δr indicates promotion after refinement, a negative value indicates demotion, and zero indicates rank preservation. This rank-shift definition is central because it converts the final fused score into an auditable movement event.

The full GLYBATOMAQ™ fusion score can be written as

$$\Lambda_{372} = \mathrm{AggP}[\Xi_{372}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{447}$$

In equation (414), preserves the upstream frontier, contributes deterministic descriptor support, contributes uncertainty-aware conservative support, contributes topology-consistency support, and penalizes unstable, incomplete, or discordant evidence. The terms and are branch reliability gates, and are fixed or logged weights.

The diagnostic penalty can be expanded as

$$\Lambda_{373} = \mathrm{AggP}[\Xi_{373}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{448}$$

where π_{DFT} penalizes weak DFT reliability, π_{QMC} penalizes excessive stochastic uncertainty or insufficient sampling, π_{topo} penalizes topology discordance, π_{prov} penalizes missing provenance, and π_{domain} penalizes descriptor-domain violations. This structure ensures that unstable evidence cannot silently improve a final rank.

The DFT proxy branch can be represented as a deterministic descriptor map,

$$\Lambda_{374} = \mathrm{AggP}[\Xi_{374}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{449}$$

where E_{el} denotes an electronic-energy proxy, $\Delta\varepsilon_{\mathrm{HL}}$ denotes a frontier-gap descriptor, μ denotes a chemical-potential-like descriptor, η denotes a hardness-like descriptor, ω denotes an electrophilicity-like descriptor, $q$denotes charge descriptors, d_{loc} denotes local contact or density descriptors, and c_{SCF} denotes convergence diagnostics. The DFT branch score is then

$$\Lambda_{375} = \mathrm{AggP}[\Xi_{375}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{450}$$

with reliability gate

$$\Lambda_{376} = \mathrm{AggP}[\Xi_{376}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{451}$$

The QMC-style branch is represented as an uncertainty-aware conservative score,

$$\Lambda_{377} = \mathrm{AggP}[\Xi_{377}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{452}$$

where $\hat{E}_{\mathrm{QMC}}$ is the estimated energy-like quantity, σ_{QMC} is the stochastic uncertainty estimate, W_{CI} is a confidence-interval width, and α, β are predeclared uncertainty penalties. The QMC reliability gate can be written as

$$\Lambda_{378} = \mathrm{AggP}[\Xi_{378}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{453}$$

where the terms denote effective sample-size adequacy, variance stability, confidence-interval acceptability, and convergence behavior.

The MQWalk branch begins from the normalized operator-overlap graph. If $K_{\mathrm{norm}}(P_j, L_i)$ denotes the normalized pocket–ligand operator overlap, the bipartite adjacency can be written as

$$\Lambda_{379} = \mathrm{AggP}[\Xi_{379}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{454}$$

The degree-normalized adjacency is

$$\Lambda_{380} = \mathrm{AggP}[\Xi_{380}(P,L)] - \lambda BD\_B - \lambda T\Delta_{\mathrm{topo}} - \lambda UCI\_QMC \tag{455}$$

where D is the diagonal degree matrix. Teleport mixing gives

$$\Lambda_{381} = \mathrm{AggP}[\Xi_{381}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{456}$$

with teleport parameter τ. The Avogadro-conditioned Laplacian can be written as

$$\Lambda_{382} = \mathrm{AggP}[\Xi_{382}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{457}$$

where N_A is used only as a deterministic numerical scale anchor. The propagation operator is

$$\Lambda_{383} = \mathrm{AggP}[\Xi_{383}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{458}$$

and the terminal node mass is

$$\Lambda_{384} = \mathrm{AggP}[\Xi_{384}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{459}$$

The MQWalk score can then be expressed as

$$\Lambda_{385} = \mathrm{AggP}[\Xi_{385}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{460}$$

where c_i denotes stationary centrality and d_i denotes topology discordance. One explicit topology-discordance form is

$$\Lambda_{386} = \mathrm{AggP}[\Xi_{386}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{461}$$

where S_{branch} summarizes non-topological support and S_{topo} summarizes graph-derived support.

The SI-derived branch-score audit demonstrates that the full fusion model does not merely reproduce the input docking order. The leading candidate retained its top position:

$$\Lambda_{387} = \mathrm{AggP}[\Xi_{387}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{462}$$

HSX-cons1 remained second:

$$\Lambda_{388} = \mathrm{AggP}[\Xi_{388}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{463}$$

However, the lower pair changed order. HSX-cons3 was promoted:

$$\Lambda_{389} = \mathrm{AggP}[\Xi_{389}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{464}$$

HSX-cons4 was demoted:

$$\Lambda_{390} = \mathrm{AggP}[\Xi_{390}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{465}$$

Equations (429)-(432) are central to the numerical interpretation. They show that the refinement layer preserved the two leading positions but changed the lower ordering after branch-level audit. Therefore,

$$\Lambda_{391} = \mathrm{AggP}[\Xi_{391}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{466}$$

and the final ranking is not a simple restatement of the docking-energy overview.

The branch-level evidence also clarifies why the ordering changed. For HSX-cons2,

$$\Lambda_{392} = \mathrm{AggP}[\Xi_{392}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{467}$$

For HSX-cons1,

$$\Lambda_{393} = \mathrm{AggP}[\Xi_{393}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{468}$$

For HSX-cons3,

$$\Lambda_{394} = \mathrm{AggP}[\Xi_{394}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{469}$$

For HSX-cons4,

$$\Lambda_{395} = \mathrm{AggP}[\Xi_{395}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{470}$$

These values explain the ablation behavior. HSX-cons4 is competitive in the Vina near-tie region, but its DFT proxy support is weak. Therefore, removal of the DFT proxy branch would weaken the explanation for its demotion. Conversely, HSX-cons3 receives relatively strong QMC-style and MQWalk support despite weak base support, explaining its promotion from fourth to third after fusion.

The leave-one-component ablation can be expressed generally as

$$\Lambda_{396} = \mathrm{AggP}[\Xi_{396}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{471}$$

where $k \in \{\mathrm{DFT}, \mathrm{QMC}, \mathrm{MQWalk}\}$, with reliability factors included where appropriate. For the diagnostic penalty ablation,

$$\Lambda_{397} = \mathrm{AggP}[\Xi_{397}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{472}$$

doi:10.20944/preprints202605.2096.v1

Equations (363) and (364) define the conceptual leave-one-component tests. They show that each branch can be removed to determine whether the final ordering depends on deterministic descriptor evidence, uncertainty-aware support, topology validation, or diagnostic penalties. The resulting interpretation is

$$\Lambda_{398} = \mathrm{AggP}[\Xi_{398}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{473}$$

Equation (440) summarizes the non-redundancy of the framework: each term addresses a different failure mode of virtual-screening interpretation.

The topology-specific findings further strengthen this conclusion. Table 7 defines the teleport-mixed transition structure of the HSX candidate graph, while Table 8 reports terminal node mass, stationary centrality, MQWalk topology score, and topology discordance. The topology outputs can be summarized as

$$\Lambda_{399} = \mathrm{AggP}[\Xi_{399}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{474}$$

and

$$\Lambda_{400} = \mathrm{AggP}[\Xi_{400}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{475}$$

Thus,

$$\Lambda_{401} = \mathrm{AggP}[\Xi_{401}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{476}$$

At the same time,

$$\Lambda_{402} = \mathrm{AggP}[\Xi_{402}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{477}$$

Equations (370) and (371) are important because they show that MQWalk is not merely a non-functional score appended to the docking winner. If MQWalk simply reinforced the top scalar candidate, HSX-cons2 would also have the strongest topology score. Instead, MQWalk identifies a topology-consistency limitation for HSX-cons2 while the full fusion still retains HSX-cons2 as the leading candidate because its base, DFT proxy, and QMC-style supports are sufficiently strong.

The topology-discordance pattern can be written as

$$\Lambda_{403} = \mathrm{AggP}[\Xi_{403}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{478}$$

whereas

$$\Lambda_{404} = \mathrm{AggP}[\Xi_{404}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{479}$$

Equations (372) and (373) show that HSX-cons2 is not universally favored across all evidence channels. Rather, its final prioritization reflects an evidence balance: strong base and descriptor/uncertainty-aware support outweigh weaker topology consistency after diagnostic penalties are applied.

The residue-resolved energetic fingerprints provide an additional interpretability layer. For ligand L_i, the residue/contact energetic vector can be written as

$$\Lambda_{405} = \mathrm{AggP}[\Xi_{405}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{480}$$

where $e_k(L_i)$ denotes the contribution of residue or contact feature k. The total interface energy is then

$$\Lambda_{406} = \mathrm{AggP}[\Xi_{406}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{481}$$

The hotspot support is

$$\Lambda_{407} = \mathrm{AggP}[\Xi_{407}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{482}$$

Two ligands may have similar total interface energies while relying on different residue/contact supports. Therefore,

$$\Lambda_{408} = \mathrm{AggP}[\Xi_{408}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{483}$$

Equation (452) explains why the rank-audit framework is more informative than a single docking or total-energy table. Similar scalar values can arise from different sparse energetic fingerprints, and those differences may have interpretive relevance for rank stability, candidate diversity, and mechanistic plausibility.

The component-wise difference between two residue/contact fingerprints can be written as

$$\Lambda_{409} = \mathrm{AggP}[\Xi_{409}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{484}$$

with magnitude

doi:10.20944/preprints202605.2096.v1

$$\Lambda_{410} = \mathrm{AggP}[\Xi_{410}(P,L)] - \lambda BD\_B - \lambda T\Delta_{topo} - \lambda UCI\_QMC \tag{485}$$

Thus, two candidates can satisfy

$$S\_fusion = \sigma(a_0+a^{T}z) \tag{486}$$

while also satisfying

$$\mathrm{Flag}(L) = 1[\Pi\_diag>\pi^{d}{}_{ia9} \text{ or } \Pi\_QG>\pi QG] \tag{487}$$

Equations (380) and (381) describe the situation in which total interface scores are similar but the energetic wiring diagrams differ substantially. In such cases, residue/contact fingerprints, branch-specific support scores, and topology diagnostics should be interpreted jointly.

The combined interpretation of Tables 6–8 and Table 10 is therefore as follows. Table 6 provides the branch-level support, diagnostic penalties, fused scores, and rank shifts. Table 7 provides the teleport-mixed transition structure used for MQWalk topology validation. Table 8 provides the terminal node mass, stationary centrality, topology score, and topology-discordance outputs. Together, these tables show that post-refinement ranking depends on three linked objects:

$$M(L) = \langle \Phi_{h\gamma p}(\varrho P),\Phi_{h\gamma p}(\varrho L)\rangle \tag{488}$$

This rank-audit object is the practical numerical output of the framework. A candidate is not promoted because one number is favorable. A candidate is promoted when the available evidence supports the movement after reliability gating, uncertainty adjustment, topology checking, and diagnostic penalization.

Accordingly, the final prioritization can be stated conservatively as

$$F\_Q(\theta) = 4[\langle \partial\theta\psi | \partial\theta\psi\rangle - |\langle \psi | \partial\theta\psi\rangle|^2] \tag{489}$$

under the supplied docking and SI-derived audit framework. However,

$$\Delta\gamma(C) = \oint C\, i\langle \psi(x) | \nabla\psi(x)\rangle\cdot dx \tag{490}$$

Equation (459) is essential for the scope of the manuscript. The evidence supports rank prioritization, internal consistency, near-tie auditing, and topology-aware interpretability. It does not establish experimental pharmacology.

The present benchmark should therefore be understood as an internal, file-derived benchmark and ablation analysis. A full publication-level virtual-screening validation would require external active/decoy or affinity datasets. The recommended external validation set can be written as

$$\Pi_{9au9e} = \lambda\gamma\ |\Delta\gamma(C)-2\pi m| \tag{491}$$

For active/decoy benchmarks, the appropriate metrics include

$$\Pi\_QG = \Pi B + \Pi_{curv} + \Pi_{9au9e} + \Pi_{path} \tag{492}$$

For affinity-ranking benchmarks, the appropriate metrics include

$$\mathrm{Cert}(L) = 1[\ \bar{K} \geq \tau,\ D\_B \leq \delta,\ S_{path} \geq \eta,\ \Pi\_QG \leq \pi_{max}\ ] \tag{493}$$

where RMSE and MAE should be used only if the model is explicitly calibrated for regression against experimental affinity values.

Statistical uncertainty should be reported as

$$S\_fusion = \sigma(a_0+a^{T}z) \tag{494}$$

where $\widehat{M}$ is the estimated benchmark metric. Pairwise comparison against docking-only and simpler baselines should use paired target-level or bootstrap-based tests,

$$\mathrm{Flag}(L) = 1[\Pi\_diag>\pi^{d}{}_{ia9} \text{ or } \Pi\_QG>\pi QG] \tag{495}$$

for $k \in \{0,1,2,\ldots,5\}$, with correction for multiple comparisons where appropriate.

Until such external validation is completed, the defensible conclusion is conservative. GLYBATOMAQ™ improves the interpretability of internal rank behavior relative to docking-only baselines by decomposing candidate movement into base support, DFT proxy descriptor evidence, QMC-style uncertainty-aware support, MQWalk topology consistency, and diagnostic penalties. HSX-cons2/CONS2 is the strongest computational hypothesis under the reported docking and SI-derived audit framework, but definitive claims of virtual-screening superiority require public data,

doi:10.20944/preprints202605.2096.v1

public code, fixed protocols, external active/decoy testing, branch ablations, and statistical comparison against simpler baselines.

4.16. Relevance to Pattern Recognition

For a Pattern Recognition readership, GLYBATOMAQ™ can be viewed as a domain-specific case of structured ranking, controlled fusion, and graph-aware validation. Recent work has emphasized ranking-based learning [56], local-kernel graph learning [57], informative graph selection [58], dynamic graph representation [59], adaptive graph propagation [60], and molecular graph contrastive learning [61]. GLYBATOMAQ™ applies related principles to post-docking molecular prioritization.

The structured-data problem is clear. A ligand candidate is not merely a vector of descriptors. It is a composite object:

$$M(L) = \langle \Phi_{h\gamma p}(\varrho P), \Phi_{h\gamma p}(\varrho L) \rangle \tag{496}$$

The ranking problem is likewise structured:

$$F\_Q(\theta) = 4[\langle \partial\theta\psi | \partial\theta\psi \rangle - |\langle \psi | \partial\theta\psi \rangle|^2] \tag{497}$$

with explanatory variables attached to the movement.

The graph problem appears through the normalized overlap kernel and MQWalk topology validation. The uncertainty problem appears through QMC and reliability gates. The interpretability problem appears through branch contributions, residue fingerprints, hotspot persistence, and diagnostic flags.

The audit-forward reporting panels in SI Appendix I Figure 3, the biology-facing optimization view in SI Appendix I Figure 4, and the rank-centric docking summaries in SI Appendix I Figure 7, SI Appendix I Figure 8, and SI Appendix I Figure 9 all reinforce this structured Pattern Recognition interpretation.

Thus, the methodological value of GLYBATOMAQ™ lies not in claiming a universal molecular score, but in showing how structured evidence can be fused, constrained, and audited in a chemically meaningful ranking workflow.

4.17. Relationship to AlphaFold, AlphaFold 3, and DiffDock

Modern upstream systems such as AlphaFold [53], AlphaFold 3 [54], and DiffDock [55] have transformed structural modeling and docking-oriented workflows. However, they do not remove the need for downstream interpretability, uncertainty handling, and rank auditing.

GLYBATOMAQ™ is compatible with these systems because it operates after structure generation, pocket definition, docking, or candidate proposal. If an upstream model supplies a receptor structure, complex hypothesis, pose ensemble, or docking-informed state, GLYBATOMAQ™ can treat that output as part of the protocol-logged input. The same audit rules still apply:

$$\Delta\gamma(C) = \oint C\, i\langle \psi(x) | \nabla\psi(x) \rangle \cdot dx \tag{498}$$

The framework therefore does not compete with AlphaFold, AlphaFold 3, or DiffDock. It addresses a downstream bottleneck: how to interpret, refine, and audit ranked molecular hypotheses once they have been generated.

This compatibility is important because modern computational biology often produces plausible structures and ranked candidate lists faster than they can be experimentally validated. GLYBATOMAQ™ contributes a mathematical-chemistry audit layer for that setting.

4.18. Limitations

The limitations of GLYBATOMAQ™ are clear and should be stated explicitly.

First, the framework is protocol-dependent. Receptor choice [2], structure selection [3,4], pocket definition [5], docking configuration [6-8,13], benchmark context [14-16], and geometry preprocessing [17] all affect the frontier subjected to refinement.

Second, representation quality depends on the adequacy of the selected cheminformatics and graph abstractions [18-25,28-32]. If the molecular representation fails to encode relevant chemistry, the operator-kernel layer cannot recover that missing information.

Third, the PSD operator geometry depends on the construction of Z. Although

$$\Pi_{9au9e} = \lambda\gamma \, |\Delta\gamma(C)\text{-}2\pi m| \tag{499}$$

guarantees positive semidefiniteness, it does not guarantee biological relevance. The quality of the operator space depends on the quality of the descriptors.

Fourth, DFT refinement depends on geometry quality, basis choices, functional choices, convergence behavior, and descriptor calibration [44-47]. The DFT branch is therefore useful only when reliability gates are applied.

Fifth, QMC refinement is computationally expensive and is used selectively [48,49]. It is not a screening-scale replacement for docking. Its role is uncertainty-aware auditing of top candidates or near-tie cases.

Sixth, MQWalk topology validation depends on the informativeness of the overlap graph [33-36,50-52]. If the graph is poorly constructed, sparse in the wrong way, or dominated by uninformative descriptors, topology validation may be weak or misleading.

Seventh, the Avogadro-conditioned scaling is a numerical regularization device, not a physical law. It improves reproducibility and boundedness but should not be overinterpreted.

Eighth, the HMC and HSX consensus structures are reference abstractions, not validated therapeutic entities and not synthesis instructions.

Ninth, the results remain in silico prioritization hypotheses. They do not establish binding affinity, efficacy, selectivity, toxicity, pharmacokinetics, or clinical performance.

These limitations are not secondary caveats. They are part of the intended conservative interpretation of the framework.

4.19. Overall Significance

Overall, GLYBATOMAQ™ is best regarded as a **triage-and-audit framework** for post-docking molecular prioritization. Docking defines the initial frontier. Operator-kernel geometry stabilizes ligand–pocket comparability. DFT provides deterministic electronic near-tie discrimination. QMC provides uncertainty-aware energetic auditing. MQWalk checks graph-topology consistency. Diagnostic penalties prevent unstable evidence from being hidden. The final result is a replayable, explainable leaderboard rather than an opaque score list.

In compact form, the framework can be summarized as

$$\Pi\_QG = \Pi B + \Pi_{curv} + \Pi_{9au9e} + \Pi_{path} \tag{500}$$

or mathematically as

$$\text{Cert}(L) = 1[\ \bar{K} \geq \tau,\ D\_B \leq \delta,\ S_{path} \geq \eta,\ \Pi\_QG \leq \pi_{max}\ ] \tag{501}$$

with rank movement reported as

$$S\_fusion = \sigma(a_0 + a^T z) \tag{502}$$

For mathematical chemistry, the contribution lies in making ligand prioritization a structured, reproducible, uncertainty-aware ranking problem. For Pattern Recognition, the contribution lies in demonstrating how kernel geometry, graph propagation, reliability-gated fusion, and interpretability diagnostics can be applied to molecular screening.

The main conclusion is therefore deliberately conservative:

$$\text{Flag}(L) = 1[\Pi\_diag > \pi^{d}{}_{ia9} \text{ or } \Pi\_QG > \pi QG] \tag{503}$$

Within that boundary, the framework provides research-level usefulness because it explains not only **which** candidates are prioritized, but **why** they move in rank, **which** evidence supports the movement, **which** uncertainty constrains it, and **whether** the movement remains compatible with the learned ligand–pocket topology.

5. Conclusions

Accordingly, the framework is positioned as a quantum-geometric decision layer for computational drug-design workflows: it preserves the rank-centric nature of virtual screening while providing a mathematically inspectable route from raw docking output to reliability-weighted, topology-aware, and uncertainty-conscious molecular prioritization.

The final interpretation remains deliberately conservative. GLYBATOMAQ™ produces auditable computational prioritization hypotheses; it does not establish experimental binding affinity, pharmacological efficacy, toxicity, selectivity, or clinical performance. Its practical value is to make post-docking ranking more transparent by connecting rank shifts to electronic descriptors, stochastic uncertainty, topology support, curvature penalties, and provenance-controlled diagnostic flags.

The integrated figure and table set supports this interpretation without reproducing the supplementary figures in the main text. Figures 10-14 and Tables 12-18 provide the quantum-geometric manifold, hypergeometric feature channels, QMC-DFT-MQWalk audit circuit, proxy-overlap generalization, topology-validation graph, and numerical diagnostic matrices. SI Appendix I Figs. 1-9 and SI Appendix II provide the supporting numerical and algorithmic context for these main-text summaries.

The manuscript formalizes each candidate as a protocol-logged molecular evidence object embedded in a normalized positive-semidefinite operator space. In that space, docking provides the fixed comparative frontier, DFT supplies deterministic electronic-structure descriptors for near-tie discrimination, QMC contributes uncertainty-aware energetic auditing, and MQWalk tests whether post-refinement promotions remain compatible with the learned operator-overlap topology.

This study presents GLYBATOMAQ™ as an original quantum-geometric rank-audit framework for GLIPR1-focused virtual screening. The central contribution is not a new absolute affinity estimator, but a reproducible mathematical procedure for explaining how a ligand moves within a ranked shortlist after electronic, stochastic, and graph-topological evidence has been integrated.

CRediT Author Contributions: Ioannis Grigoriadis: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Visualization, Writing—Original Draft, Writing—Review & Editing, Project Administration, Resources.

Data Availability Statement: The datasets produced and analyzed in this study, including docking results, rank-shift diagnostics, and operator-kernel artifacts, are not in public repositories. However, the underlying computational outputs—candidate rank logs, topology graphs, DFT/QMC refinements, and fusion scores—are available from the corresponding author upon reasonable request, subject to institutional or confidentiality constraints.

Code Availability Statement: The custom code, including operator-kernel construction, DFT/QMC integration, and MQWalk validation, is not publicly available. It can be requested from the corresponding author, pending any third-party licensing limitations.

Funding: This work was funded internally. No external grants or funding agencies supported this research.

Declaration of Interest: The author, Ioannis Grigoriadis, declares financial and intellectual property interests related to Biogenea™ and its R&D operations. The work was conducted without external sponsor influence. Aside from these affiliations, no other competing interests are declared.

Acknowledgements: The author thanks Biogenea™ and affiliated R&D units for their support, infrastructure, and computational resources.

Generative AI and Image Integrity Disclosure: Generative AI disclosure (text): Generative AI tools assisted with language refinement, section structuring, and clarity enhancement. All scientific decisions, interpretations, and conclusions were independently made by the author.

Generative AI for images: AI-assisted visuals were used to depict the GLYBATOMAQ™ computational workflow—including topology diagnostics, rank-shift schematics, and audit panels—and are not experimental data. Inputs (e.g., similarity matrices, transport distributions) are documented, and the author ensures all images represent the method transparently.

doi:10.20944/preprints202605.2096.v1

Table 11. Symbols and abbreviations used in the GLYBATOMAQ™ framework.

Symbol / Abbreviation	Meaning / Definition
GLYBATOMAQ™	GLYBATOMAQ pipeline with late-stage quantum-enriched refinement and topology validation layer
P	Protein pocket (under a fixed docking protocol)
L	Ligand / candidate molecule
(P, L)	Pocket–ligand candidate pair
$\mathcal{L}_P$	Full set of docked ligands for pocket P
$\mathcal{L}_{\text{top}}(P)$	Top-K shortlist of ligands for pocket P by base GLYBATOMAQ score
$\mathcal{L}_{\text{top}}^{\text{cons}}$	Consensus shortlist pooled across pockets
$\mathcal{P}(T)$	Set of pockets considered for target T
$D(P, L)$	Docking oracle score (lower is better, by manuscript convention)
$S_\theta(P, L)$	Base GLYBATOMAQ score prior to quantum enrichment
$S_{\text{DFT}}(P, L)$	DFT branch refinement score
$S_{\text{QMC}}(P, L)$	QMC branch refinement score (uncertainty-aware)
$S_{\text{MQW}}(P, L)$	MQWalk topology validation / propagation score
$S_{\text{final}}^{(\quad + \quad)}(P, L)$	Final fused quantum-enriched GLYBATOMAQ™ score
$S_{\text{cons}}^{(\quad + \quad)}(L)$	Consensus score across pockets after quantum enrichment
$S_{\text{cons}}(L)$	Consensus score before quantum enrichment
ρ_P, ρ_L	PSD operator representations of pocket and ligand features
$K(P, L)$	Unnormalized trace-overlap kernel, $K = \text{Tr}(\rho_P \rho_L)$
$\bar{K}(P, L)$	Normalized trace-overlap kernel
ε	Small positive numerical stabilizer
N_A	Avogadro constant used here as deterministic magnitude conditioner
$s(\alpha)$	Avogadro-conditioning factor, $s(\alpha) = \exp(-\alpha \log N_A)$
α	Avogadro-conditioning exponent / scaling hyperparameter
$f_k, z_k, \tilde{f}_k, \hat{f}_k$	Raw, robust-normalized, Avogadro-conditioned, and bounded feature channel values
MAD	Median absolute deviation
$\text{Conf}(L)$	Conformer ensemble of ligand L
χ_c	c-th conformer of ligand L
$\chi^\star(P, L)$	Selected conformer/geometry for quantum refinement
$O(P, L)$	Unified candidate object passed to DFT/QMC/MQWalk branches
DFT	Density functional theory
SCF	Self-consistent field
$\psi_i(r)$	Kohn–Sham orbital
ϵ_i	Kohn–Sham orbital energy eigenvalue
$V_{\text{eff}}(r),\ V_{xc}(r)$	Effective and exchange-correlation Kohn–Sham potentials
$E[\rho]$	DFT energy functional
E_{DFT}	DFT total electronic energy (branch descriptor)
Δ_{HL}	Frontier-orbital gap proxy (e.g., HOMO–LUMO gap)
μ	Dipole-related summary (if included)
q_{stats}	Charge distribution summary statistics
SCFdiag	SCF convergence diagnostics summary
$\Phi_{\text{DFT}}(P, L)$	DFT descriptor vector
β_{DFT}	Coefficient vector for linear DFT branch score
QMC	Quantum Monte Carlo (VMC/DMC-style)
VMC / DMC	Variational / Diffusion Monte Carlo
Ψ, Ψ_0	Trial wave function; initial trial wave function
$\hat{H}$	Hamiltonian operator
$\hat{E}_{\text{QMC}}(P, L)$	Estimated QMC energy-like quantity
σ_{QMC}	QMC uncertainty summary (SD/SE)

Symbol / Abbreviation	Meaning / Definition
CIwidth	Confidence interval width for QMC estimate
ESS	Effective sample size
$\Phi_{\mathrm{QMC}}(P,L)$	QMC descriptor vector
λ_σ,λ_c	Penalty weights for QMC uncertainty terms
$r_{\mathrm{DFT}}(P,L)$	DFT branch reliability factor in $[0,1]$
$r_{\mathrm{QMC}}(P,L)$	QMC branch reliability factor in $[0,1]$
MQWalk	Avogadro-conditioned topology walk/teleport validation module
W	Bipartite adjacency matrix built from $\bar{K}$
D (graph)	Degree matrix associated with W (distinct from docking score $D(P,L)$)
A	Degree-normalized adjacency matrix, $A = D^{-1/2}WD^{-1/2}$
A_γ	Teleportation-mixed normalized adjacency matrix
γ	Teleportation mixing coefficient
$\tilde{L}$	Avogadro-conditioned graph Laplacian
$U(t)$	MQWalk propagation operator, $U(t) = \exp(-it\tilde{L})$
t	Walk time / propagation hyperparameter
$\mathrm{MQWscore}(P,L)$	Ligand-side propagation score from MQWalk
$\Delta_{\mathrm{topo}}(P,L)$	Topology disagreement penalty
$\lambda_{\mathrm{walk}},\lambda_{\mathrm{disc}}$	MQWalk mixing and topology-disagreement weights
$R_0(P,L),\ R_+(P,L)$	Pre- and post-refinement ranks
$\Delta R(P,L)$	Rank shift $= R_0 - R_+$ (positive = promotion)
$\Pi_{\mathrm{diag}}(P,L)$	Diagnostic penalty term in final fusion
$w_0, w_{\mathrm{DFT}}, w_{\mathrm{QMC}}, w_{\mathrm{MQW}}$	Fusion weights in final score
Θ	Set of fitted/learnable refinement/fusion parameters
$\mathcal{P}(P)$	Docking-induced preference-pair set within pocket P
m	Margin parameter in pairwise ranking loss
w_{ij}	Pairwise ranking weight (e.g., hardness-based)
$\mathcal{L}(\Theta;P)$	Pocket-specific training/fit loss for refinement layer
$\widehat{D}^{(\quad+\quad)}(P,L)$	Optional calibrated projection to docking-like reporting scale
$\mathrm{Overlap@k}(P)$	Top-k overlap between pre/post rankings for pocket P
GLYBATOMAQ_HMC	GLYBATOMAQ Hyper Mega Core conceptual modular scaffold
$R1, R2$	Mapped HMC attachment handles for sulfonate-cap and piperazine-tail substitution
X1–X5	HMC motif-emphasis computational variant families
Markush/YAML	Enumeration schema describing mapped core, reactions, and fragment catalogs
Hierarchical consensus hyper-core	Family-level scaffold architecture + shared pharmacophore grammar representation (not literal atom-level fused scaffold)

References

1. Berman HM, Westbrook J, Feng Z, Gilliland TN, Bhat H, Weissig J, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235-242.
2. Asojo OA, Koski RA, Bonafé N. Structural studies of human glioma pathogenesis-related protein 1. *Acta Crystallogr D Biol Crystallogr.* 2011;67:847-855. doi:10.1107/S0907444911028198.
3. RCSB Protein Data Bank. PDB ID 3Q2U: Structure of human glioma pathogenesis-related protein 1 [Internet]. Accessed 2026.
4. RCSB Protein Data Bank. PDB ID 3Q2R: Crystal structure of sGLIPR1 soaked with zinc chloride [Internet]. Accessed 2026.
5. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics.* 2009;10:168.
6. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455-461.

7. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30(16):2785-2791.
8. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem.* 2004;47(7):1739-1749.
9. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov.* 2004;3(11):935-949.
10. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, et al. A critical assessment of docking programs and scoring functions. *J Med Chem.* 2006;49(20):5912-5931.
11. Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. *Biophys Rev.* 2017;9(2):91-102.
12. Guedes IA, Pereira FSS, Dardenne LE. Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges. *Front Pharmacol.* 2018;9:1082.
13. Guedes IA, Pereira da Silva MM, Galheigo M, Krempser E, de Magalhães CS, Correa Barbosa HJ, Dardenne LE. DockThor-VS: a free platform for receptor-ligand virtual screening. *J Mol Biol.* 2024;436:168548.
14. Wang R, Fang X, Lu Y, Yang CY, Wang S. The PDBbind database: methodologies and updates. *J Med Chem*.2005;48(12):4111-4119.
15. Huang N, Shoichet BK, Irwin JJ. Benchmarking sets for molecular docking. *J Med Chem.* 2006;49(23):6789-6801.
16. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem.* 2012;55(14):6582-6594.
17. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem.* 1996;17(5-6):490-519.
18. Landrum G. RDKit: Open-source cheminformatics [Internet]. 2006-.
19. Morgan HL. The generation of a unique machine description for chemical structures-A technique developed at Chemical Abstracts Service. *J Chem Doc.* 1965;5(2):107-113.
20. Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model.* 2010;50(5):742-754.
21. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci.* 1988;28(1):31-36.
22. Krenn M, Häse F, Nigam A, Friederich P, Aspuru-Guzik A. Self-referencing embedded strings (SELFIES): a 100% robust molecular string representation. *Mach Learn Sci Technol.* 2020;1(4):045024.
23. Kusner MJ, Paige B, Hernández-Lobato JM. Grammar variational autoencoder. In: *Proc ICML*; 2017. p. 1945-1954.
24. Jin W, Barzilay R, Jaakkola T. Junction tree variational autoencoder for molecular graph generation. In: *Proc ICML*; 2018. p. 2323-2332.
25. You J, Liu B, Ying Z, Pande V, Leskovec J. Graph convolutional policy network for goal-directed molecular graph generation. In: *Adv Neural Inf Process Syst*; 2018. p. 6410-6421.
26. De Cao N, Kipf T. MolGAN: An implicit generative model for small molecular graphs. arXiv:1805.11973; 2018.
27. Shi C, Luo S, Xu M, Tang J. GraphAF: a flow-based autoregressive model for molecular graph generation. In: *Proc ICLR*; 2020.
28. Segler MHS, Kogej T, Tyrchan C, Waller MP. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent Sci.* 2018;4(1):120-131.
29. Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent Sci.* 2018;4(2):268-276.
30. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. In: *Proc ICML*; 2017. p. 1263-1272.
31. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: *Proc ICLR*; 2017.
32. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. In: *Adv Neural Inf Process Syst*; 2017. p. 5998-6008.

33. Schölkopf B, Smola AJ. *Learning with kernels: support vector machines, regularization, optimization, and beyond.* Cambridge (MA): MIT Press; 2002.
34. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20(3):273-297.
35. Rasmussen CE, Williams CKI. *Gaussian processes for machine learning.* Cambridge (MA): MIT Press; 2006.
36. Bhatia R. *Positive definite matrices.* Princeton (NJ): Princeton University Press; 2007.
37. Mason JC, Handscomb DC. *Chebyshev polynomials.* Boca Raton (FL): CRC Press; 2003.
38. Trefethen LN. *Approximation theory and approximation practice.* Philadelphia (PA): SIAM; 2013.
39. Abramowitz M, Stegun IA, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* New York: Dover; 1965.
40. Olver FWJ, Lozier DW, Boisvert RF, Clark CW, editors. *NIST handbook of mathematical functions.* Cambridge: Cambridge University Press; 2010.
41. Andrews GE, Askey R, Roy R. *Special functions.* Cambridge: Cambridge University Press; 1999.
42. Koekoek R, Lesky PA, Swarttouw RF. *Hypergeometric orthogonal polynomials and their q-analogues.* Berlin: Springer; 2010.
43. Mohr PJ, Newell DB, Taylor BN. CODATA recommended values of the fundamental physical constants: 2018. *Rev Mod Phys.* 2021;93(2):025010.
44. Hohenberg P, Kohn W. Inhomogeneous electron gas. *Phys Rev.* 1964;136(3B):B864-B871.
45. Kohn W, Sham LJ. Self-consistent equations including exchange and correlation effects. *Phys Rev.*1965;140(4A):A1133-A1138.
46. Parr RG, Yang W. *Density-functional theory of atoms and molecules.* New York: Oxford University Press; 1989.
47. Martin RM, Reining L, Ceperley DM. *Interacting electrons: theory and computational approaches.* Cambridge: Cambridge University Press; 2016.
48. Foulkes WMC, Mitas L, Needs RJ, Rajagopal G. Quantum Monte Carlo simulations of solids. *Rev Mod Phys.*2001;73(1):33-83.
49. Needs RJ, Towler MD, Drummond ND, López Ríos P. Continuum variational and diffusion quantum Monte Carlo calculations. *J Phys Condens Matter.* 2010;22(2):023201.
50. Nielsen MA, Chuang IL. *Quantum computation and quantum information.* 10th anniversary ed. Cambridge: Cambridge University Press; 2010.
51. Arndt M, Juffmann T, Vedral V. Quantum physics meets biology. *HFSP J.* 2009;3(6):386-400.
52. Nunn AVW, Guy GW, Bell JD. Thermodynamics and inflammation: insights into quantum biology and ageing. *Quantum Rep.* 2022;4(1):47-74.
53. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7887):583-589.
54. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature.* 2024;630:493-500.
55. Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T. DiffDock: diffusion steps, twists, and turns for molecular docking. In: *Proc ICLR*; 2023.
56. Li J, Zhu X, Zhang W, Wang J. A ranking-based problem transformation method for weakly supervised multi-label learning. Pattern Recognit. 2024;153:110505. doi:10.1016/j.patcog.2024.110505.
57. Liu Z, Huang S, Jin W, Mu Y. Local kernels based graph learning for multiple kernel clustering. Pattern Recognit. 2024;150:110300. doi:10.1016/j.patcog.2024.110300.
58. Ju W, Mao Z, Qiao Z, Qin Y, Ren P, Feng X, et al. Focus on informative graphs! Semi-supervised active learning for graph-level classification. Pattern Recognit. 2024;153:110567. doi:10.1016/j.patcog.2024.110567.
59. Yao HY, Zhang CY, Yao ZL, Chen CLP, Hu J. A recurrent graph neural network for inductive representation learning on dynamic graphs. Pattern Recognit. 2024;154:110577. doi:10.1016/j.patcog.2024.110577.
60. Chen W, Yan W, Wang W. Adaptive propagation deep graph neural networks. Pattern Recognit. 2024;154:110607. doi:10.1016/j.patcog.2024.110607.
61. Chen X, Li S, Liu R, Shi B, Xu K. Molecular graph contrastive learning with line graph. Pattern Recognit. 2025;162:111380. doi:10.1016/j.patcog.2025.111380.

62. Provost JP, Vallee G. Riemannian structure on manifolds of quantum states. Commun Math Phys. 1980;76:289-301.
63. Berry MV. Quantal phase factors accompanying adiabatic changes. Proc R Soc Lond A. 1984;392:45-57.
64. Braunstein SL, Caves CM. Statistical distance and the geometry of quantum states. Phys Rev Lett. 1994;72(22):3439-3443.
65. Bures D. An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras. Trans Am Math Soc. 1969;135:199-212.
66. Uhlmann A. The transition probability in the state space of a *-algebra. Rep Math Phys. 1976;9(2):273-279.
67. Bengtsson I, Zyczkowski K. Geometry of Quantum States: An Introduction to Quantum Entanglement. Cambridge: Cambridge University Press; 2006.
68. Petz D. Monotone metrics on matrix spaces. Linear Algebra Appl. 1996;244:81-96.
69. Amari S. Information Geometry and Its Applications. Tokyo: Springer; 2016.
70. Kempe J. Quantum random walks: an introductory overview. Contemp Phys. 2003;44(4):307-327.