Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

CO-DESIGNING PROTEINS, LIGANDS, BINDING INTERFACES, AND MOLECULAR ASSEMBLIES THROUGH CONSTRAINT-AWARE GENERATIVE DIFFUSION

George Wilson[1]*, Chloe Bennett[2], Ethan Wright[1], Jack Turner[3]

1. ***Department of Co-Design and Generative Diffusion for Molecular Assemblies, Faculty of Pharmacy, University of Dundee, Dundee, United Kingdom.***
2. ***Department of Constraint-Aware Protein and Ligand Design, Faculty of Pharmacy, Newcastle University, Newcastle, United Kingdom.***
3. ***Department of Binding Interface and Assembly Design, Faculty of Pharmacy, University of Aberdeen, Aberdeen, United Kingdom.***

ARTICLE INFO

Received:
28 August 2025
Received in revised form:
10 December 2025
Accepted:
14 December 2025
Available online:
28 December 2025

Keywords: Constraint-aware diffusion, Molecular co-design, Protein–ligand interfaces, Biomolecular assemblies, Equivariant generative modeling, Synthesizability

ABSTRACT

Generative molecular modeling increasingly supports the design of ligands, proteins, binding partners, and higher-order biomolecular systems, yet most architectures still optimize these objects as partially separated tasks. This separation is problematic because pharmaceutical function often emerges from the geometry, chemistry, conformational compatibility, cooperativity, and realizability of an interface or assembly rather than from the apparent quality of either component alone. This article develops a proposed constraint-aware generative diffusion architecture for the joint co-design of proteins, ligands, binding interfaces, and molecular assemblies. The approach treats the interacting molecular system as the primary generative object and organizes it through a typed joint molecular scene, a coupled equivariant denoising process, explicit geometric, energetic, functional, and synthetic constraint channels, a constraint-arbitration layer, and a multidimensional validation boundary. The synthesis argues that no single docking, affinity, validity, novelty, confidence, or generative score can establish successful co-design. Instead, candidate systems require separable assessment of molecular structure, cross-component interaction quality, assembly organization, uncertainty, diversity, generalization, synthesizability, expression or realization feasibility, and prospective experimental behavior. The proposed architecture also distinguishes hard requirements from approximate guidance, model confidence from calibrated uncertainty, structural plausibility from energetic or mechanistic support, and computational success from pharmaceutical usefulness. Important limitations include incomplete representation of solvent, dynamics, protonation, alternative biological states, synthesis and expression constraints, negative data, and distribution shift. The contribution is therefore conceptual and methodological rather than an implemented or validated system. It is intended to clarify what molecular co-design must represent, how competing constraints should be exposed, and what evidence is required before generated complexes can inform consequential therapeutic decisions.

To Cite This Article: Wilson G, Bennett Ch, Wright E, Turner J. Co-Designing Proteins, Ligands, Binding Interfaces, and Molecular Assemblies through Constraint-Aware Generative Diffusion. Pharmacophore. 2025;16(6):109-19. https://doi.org/10.51847/YYA2shCYJ9

Introduction

Generative molecular design reframes chemical discovery as an inverse problem in which candidate structures are produced in response to specified objectives, representations, and search constraints. Its pharmaceutical value, however, depends on more than the ability to sample novel molecular graphs or plausible protein conformations. Generation must be connected to evaluation, chemical or biological realization, and an explicit decision context, because an optimized model output may remain irrelevant to the experimental or therapeutic problem for which it was produced [1-3]. This distinction has become increasingly important as generative systems move from isolated small-molecule design toward protein engineering, three-dimensional structure generation, induced interactions, and multicomponent therapeutic systems.

Many current workflows continue to divide the molecular problem into sequential tasks. A ligand may first be generated and subsequently docked into a fixed receptor; a protein backbone may be generated before sequence assignment; a linker may be

Corresponding Author: George Wilson; Department of Co-Design and Generative Diffusion for Molecular Assemblies, Faculty of Pharmacy, University of Dundee, Dundee, United Kingdom. E-mail: g.wilson@dundee.ac.uk

optimized after the identities of its connected components have been selected; or an assembly may be inferred from independently modeled subunits. Such decompositions are computationally convenient and can be scientifically useful, but they also create a structural assumption: that the components can be optimized separately without losing the interaction biology that emerges only when they are considered together. This assumption is particularly fragile for flexible binding sites, induced-fit systems, protein–protein interfaces, molecular glues, targeted degraders, allosteric modulators, multivalent constructs, and designed assemblies in which one component changes the feasible state of another.

Recent advances in all-atom biomolecular modeling demonstrate that proteins, nucleic acids, small molecules, ions, and modified residues can be represented within a shared complex-prediction setting [4]. This development provides an important representational precedent but does not itself resolve the generative co-design problem. Predicting a likely structure from specified molecular identities differs from generating those identities, geometries, sequences, interfaces, and assembly relations under competing pharmaceutical constraints. A joint predictor can estimate a complex while leaving unanswered which molecular variables should be editable, how constraints should be enforced during generation, how incompatible requirements should be exposed, and what evidence is needed to determine whether a generated complex is chemically, physically, functionally, and synthetically credible.

This article therefore proposes a Constraint-Aware Joint Molecular Diffusion Architecture in which the primary generative object is an interacting molecular system rather than an isolated ligand or protein. The architecture comprises a typed joint molecular scene, coupled equivariant denoising, entity-specific and cross-entity state updates, a constraint compiler, a constraint-arbitration layer, an uncertainty record, and a multidimensional validation gate. Its central proposition is that proteins, ligands, interfaces, and assemblies should be co-conditioned while retaining distinct chemical and biological identities, and that success should be expressed as an auditable vector of partially independent evidence dimensions rather than a universal scalar score. The architecture is presented as an original conceptual and methodological contribution requiring future implementation and prospective validation; it is not offered as a validated predictor, experimental platform, clinical recommendation, or operational decision tool.

Why Isolated Ligand or Protein Generation Misses Interface Biology

A binding interface is not merely the spatial overlap between two otherwise independent molecular objects. It is a relational state produced by cross-component geometry, surface complementarity, local chemical environments, conformational compatibility, and, in some systems, cooperative contacts involving additional partners. Protein–protein interaction prediction improves when information from both partners is considered, geometric learning from molecular surfaces can identify complementary interaction fingerprints, and structural studies of targeted degradation show that ligand-induced higher-order contacts can reorganize recognition within a ternary complex [5-7]. Together, these findings support an interface-centered interpretation: properties relevant to binding and function may reside in relationships among components and cannot always be reconstructed reliably after each component has been optimized alone.

Isolated ligand generation can fail even when the generated molecule satisfies formal chemical validity, drug-likeness filters, or a target-conditioned score. A ligand graph does not uniquely determine its biologically relevant protonation state, tautomer, conformation, orientation, solvation pattern, or effect on the receptor ensemble. When the receptor is treated as rigid, ligand generation may exploit a pocket state that is geometrically available in one structure but energetically inaccessible or biologically irrelevant in the required context. Conversely, allowing unrestricted receptor movement can create an equally problematic form of accommodation in which the protein deforms unrealistically to accept the ligand. The co-design problem is therefore not solved simply by making both objects flexible; their permitted changes must be constrained by molecular identity, conformational evidence, functional state, and the intended mechanism.

Isolated protein generation creates a corresponding failure mode. A designed backbone can display acceptable local geometry and may even receive a plausible sequence, yet the intended partner can expose requirements that were absent from the protein-only objective. Side-chain packing, electrostatic preorganization, desolvation, shape complementarity, accessibility, and competing interaction states may alter whether the proposed surface functions as intended. These dependencies become more pronounced in ternary and higher-order systems. Studies of BTK-directed degraders, for example, show that cooperativity within a target–degrader–ligase complex can influence productive molecular behavior, making binary affinity an incomplete description of the system [8]. The implication is not that every therapeutic design requires simultaneous unrestricted generation of all components, but that fixed and editable components must be declared explicitly and evaluated within the same relational state.

The proposed architecture treats this relational state as the unit of co-design and divides interface biology into several auditable evidence domains. These include entity validity, conformational state, interface geometry, chemical complementarity, cooperative organization, functional constraints, negative-design requirements, and realization feasibility. The division is necessary because success in one domain can conceal failure in another: a highly complementary surface may be energetically implausible, a favorable predicted complex may require an inaccessible ligand, or a stable binary interaction may prevent formation of the desired ternary assembly. **Table 1** organizes the evidence, constructs, and interpretation boundaries needed to develop why isolated ligand or protein generation misses interface biology within the article's central argument.

Table 1. Evidence domains, core questions, scientific requirements, and interpretation boundaries for why isolated ligand or protein generation misses interface biology

Construct or evidence domain	Core question	Scientific basis	Role in the article	Failure or overclaiming risk	Boundary statement
Molecular component identity	Are the protein, ligand, and other components chemically and biologically defined?	Molecular identity determines permissible atoms, residues, bonds, modifications, charges, stereochemistry, and component roles.	Prevents the joint representation from collapsing distinct entities into an undifferentiated geometric object.	A geometrically coherent output may contain chemically invalid or incorrectly assigned components.	Correct component identity does not establish a correct interaction.
Protein conformational state	Is the represented protein state relevant to the intended binding or functional context?	Proteins can occupy alternative conformations with different pockets, surfaces, and interaction competencies.	Makes receptor or partner state an explicit input, editable variable, or uncertainty source.	A single available structure may be treated as the complete biological ensemble.	Structural availability does not establish state suitability.
Ligand conformational and chemical state	Are conformation, stereochemistry, protonation, tautomerism, and charge treated consistently?	Small-molecule interaction geometry depends on coupled chemical and conformational variables.	Links graph generation to three-dimensional and context-sensitive ligand representation.	Formal graph validity may be mistaken for a biologically relevant ligand state.	Chemical validity does not establish binding competence or synthesizability.
Binding-interface geometry	Are distances, orientations, contacts, cavities, and steric relations jointly plausible?	Interfaces arise from cross-component geometric complementarity rather than either component alone.	Establishes the interface as a first-class generative object.	Contact-rich structures may contain collisions, strain, or implausible orientation.	Geometric complementarity does not establish favorable binding energy.
Surface chemistry	Are hydrophobic, polar, electrostatic, hydrogen-bonding, and solvent-facing features compatible?	Local chemical environments influence recognition, desolvation, and specificity.	Adds chemically typed relations to geometric interface edges.	Simplified contact labels may be interpreted as a complete physical model.	Chemical-pattern compatibility remains an approximation until independently tested.
Conformational coupling	Does alteration of one component induce feasible changes in another?	Binding can involve induced fit, conformational selection, or coupled rearrangement.	Requires coordinated rather than independent state updates.	Flexible components may deform unrealistically to satisfy the model objective.	Predicted accommodation does not establish that the state is accessible or populated.
Cooperative higher-order recognition	Does a third component create or disrupt contacts that are absent in binary systems?	Ternary and multicomponent complexes can display emergent interfaces and cooperativity.	Supports explicit representation of assembly-level and higher-order relations.	Binary affinity may be used as a proxy for ternary or assembly behavior.	Pairwise interaction quality does not determine higher-order function.
Assembly identity and topology	Are component number, stoichiometry, orientation, symmetry, and interaction roles coherent?	Higher-order function depends on organization across several molecular entities.	Extends co-design beyond one protein–ligand pair.	A static assembly model may be mistaken for evidence of formation, persistence, or homogeneity.	Predicted topology does not establish assembly kinetics or stability.
Functional mechanism	Does the interface satisfy a mechanism-relevant requirement rather than only a structural objective?	Binding, catalysis, recruitment, inhibition, and degradation require different molecular relationships.	Connects generated structure to a declared functional hypothesis.	Structural plausibility may be presented as mechanistic explanation or causality.	A generated arrangement can suggest a mechanism but cannot confirm it.
Negative design and selectivity	Are unwanted interactions, states,	Desired binding can coexist with off-	Prevents optimization	Incomplete negative sets can	Computational negative design does

	homologs, or assembly routes explicitly considered?	target or competing interactions.	against the intended complex alone.	create false confidence in selectivity.	not establish biological selectivity.
Energetic plausibility	Are proposed contacts and conformational changes compatible with independent physical assessment?	Stable interaction depends on more than static geometry and learned scores.	Creates an evaluation channel separate from geometric generation.	A model score may be interpreted as measured affinity or free energy.	Energetic predictions are method- and context-dependent estimates.
Synthetic and biological realization	Can ligand components be synthesized and protein components be expressed, folded, and assembled?	Molecular candidates must be physically produced before their proposed behavior can be tested.	Connects generation to actionable experimental hypotheses.	Computational accessibility scores may be treated as proof of practical realization.	Predicted feasibility does not establish synthesis yield, expression, manufacturability, or therapeutic value.
Uncertainty and applicability	How reliable is each component, interface, constraint, and validation claim in the relevant domain?	Confidence can vary across structures, entities, targets, and distributional settings.	Preserves uncertainty as part of the candidate record rather than hiding it in an aggregate rank.	High model confidence may be confused with calibrated reliability.	Confidence is not evidence of correctness and may fail under distribution shift.

Representing Proteins, Ligands, Interfaces, and Assemblies Jointly

A joint molecular representation must preserve common geometric relationships without erasing the different scientific meanings of proteins, ligands, interfaces, and assemblies. Protein structure prediction has shown the value of integrating residue-level, pairwise, and coordinate information across multiple scales [9]. For co-design, this principle must be extended beyond a protein-only state. The proposed Joint Molecular Scene contains typed molecular entities positioned within a shared coordinate system, with each entity retaining its own permissible variables. Protein states may include sequence identities, residue frames, backbone and side-chain coordinates, covalent modifications, fixed motifs, and editable regions. Ligand states may include atoms, bonds, stereochemistry, protonation or tautomer alternatives, conformers, fragments, and attachment points. Other components, including cofactors, ions, nucleic acids, or additional proteins, can be represented through corresponding typed schemas rather than forced into protein or ligand categories.

The representation also requires communication among complementary information tracks. Three-track protein modeling has demonstrated that sequence, pairwise relationships, and coordinates can exchange information iteratively rather than being processed as isolated descriptions [10]. In the proposed architecture, a related principle operates across entity and relational tracks. An entity track preserves molecular composition and intramolecular connectivity; a geometric track records coordinates, frames, distances, and orientations; an interface track represents cross-component contacts and compatibility; and an assembly track records component identities, stoichiometry, symmetry, and higher-order topology. Information can pass among these tracks, but each remains inspectable so that a favorable assembly-level update cannot silently overwrite an invalid covalent structure or an implausible local conformation.

All-atom biomolecular modeling provides evidence that heterogeneous molecular types can be handled within a common representation, including proteins, nucleic acids, small molecules, metals, and covalent modifications [11]. The proposed scene builds on this representational possibility while introducing a stricter distinction between shared geometry and shared semantics. Two atoms may occupy the same coordinate space but have different permissible transformations depending on whether they belong to a protein backbone, a ligand ring, a metal-coordination environment, or a fixed catalytic motif. Similarly, a contact edge should identify not only distance but also the entities involved, the directionality or orientation of the interaction, whether the relation is required or prohibited, and the uncertainty of the assigned state. This typing is essential for constraint-aware generation because constraints must act on scientifically meaningful variables rather than on coordinates alone.

Assembly-level representation extends beyond the union of pairwise interfaces. Computed models of multicomponent protein complexes illustrate the need to reason across interchain relationships and higher-order organization [12]. A co-design architecture must therefore record component number, identity, role, stoichiometry, symmetry, connectivity, and alternative assembly hypotheses. Pairwise contact maps can remain useful, but they may not encode whether several individually plausible interfaces can coexist without steric conflict, geometric frustration, or incompatible conformational demands. The proposed Joint Molecular Scene consequently uses a hierarchical structure: atom and residue variables define components; cross-entity relations define interfaces; and assembly variables define the organization of the full molecular system. This representation does not claim to capture all solvent, kinetic, thermodynamic, cellular, or environmental determinants. Its purpose is narrower: to establish an auditable generative state in which interacting components can be co-conditioned without treating interface and assembly biology as an afterthought.

Constraint-Aware Diffusion Architecture for Molecular Co-Design

The proposed architecture uses diffusion not as an unconstrained generator of molecular objects but as a controlled process for transforming a noisy or partially specified joint molecular scene into a candidate interacting system. Conditional protein diffusion demonstrates that structural motifs, functional regions, or target surfaces can guide the generation of compatible protein backbones [13]. In the proposed extension, the denoising state includes protein coordinates, ligand atoms and bonds, interface relations, and assembly variables simultaneously. Each reverse step can revise only the variables declared editable, while fixed catalytic residues, conserved domains, ligand fragments, attachment points, target conformations, or assembly components remain protected by explicit masks. This arrangement permits several design modes within one architecture, including ligand generation around a fixed protein, protein design around a fixed ligand, reciprocal local refinement, linker generation between anchored components, and constrained construction of higher-order assemblies.

Sequence and structure are related but non-equivalent design variables. ProteinMPNN shows that amino-acid identities can be generated conditional on an existing backbone, providing a practical basis for separating backbone formation from sequence assignment [14]. The proposed architecture retains this separation while allowing iterative communication between the two processes. A backbone update can change the local sequence requirements of an interface, whereas a sequence proposal can reveal steric, electrostatic, folding, or expression conflicts that require structural revision. The sequence–structure co-design head therefore produces conditional proposals rather than final decisions. Conserved motifs, disulfide patterns, catalytic residues, post-translational modification sites, and forbidden sequence regions can be encoded as hard constraints, while developability-related preferences can enter as softer guidance requiring later verification.

For ligand-containing systems, the denoiser must preserve equivariance while respecting chemical identity. Equivariant diffusion has been applied to structure-based ligand generation within protein environments, showing how three-dimensional molecular proposals can be conditioned on target geometry without dependence on an arbitrary coordinate frame [15]. The proposed coupled denoiser extends that logic by allowing limited, explicitly bounded adjustment of both the ligand and selected protein regions. Ligand coordinates, atom identities, bond states, stereochemistry, and fragment membership remain distinguishable from protein residue and backbone variables. Coordinate updates can be shared across the scene, but chemical transformations remain governed by entity-specific rules. This distinction prevents geometric compatibility from being achieved through chemically meaningless operations and permits uncertainty about protonation, tautomerism, side-chain placement, or alternative conformations to remain visible.

Constraint-aware linker generation illustrates how fixed and generated molecular regions can coexist within a diffusion process. Conditional linker models can preserve molecular endpoints while generating connecting structures and conformations [16]. Within the proposed architecture, this principle is generalized into a constraint compiler that translates scientific requirements into fixed masks, geometric anchors, conditional features, soft guidance terms, prohibited relations, or post-generation tests. A separate arbitration layer identifies incompatible requirements rather than concealing them inside a weighted objective. Candidate states can advance, undergo local revision, be resampled, or be rejected according to which constraint failed and how confidently that failure can be assessed. **Figure 1** presents the protein–ligand co-design diffusion chamber, showing how the article's key components and boundaries are connected within constraint-aware diffusion architecture for molecular co-design.

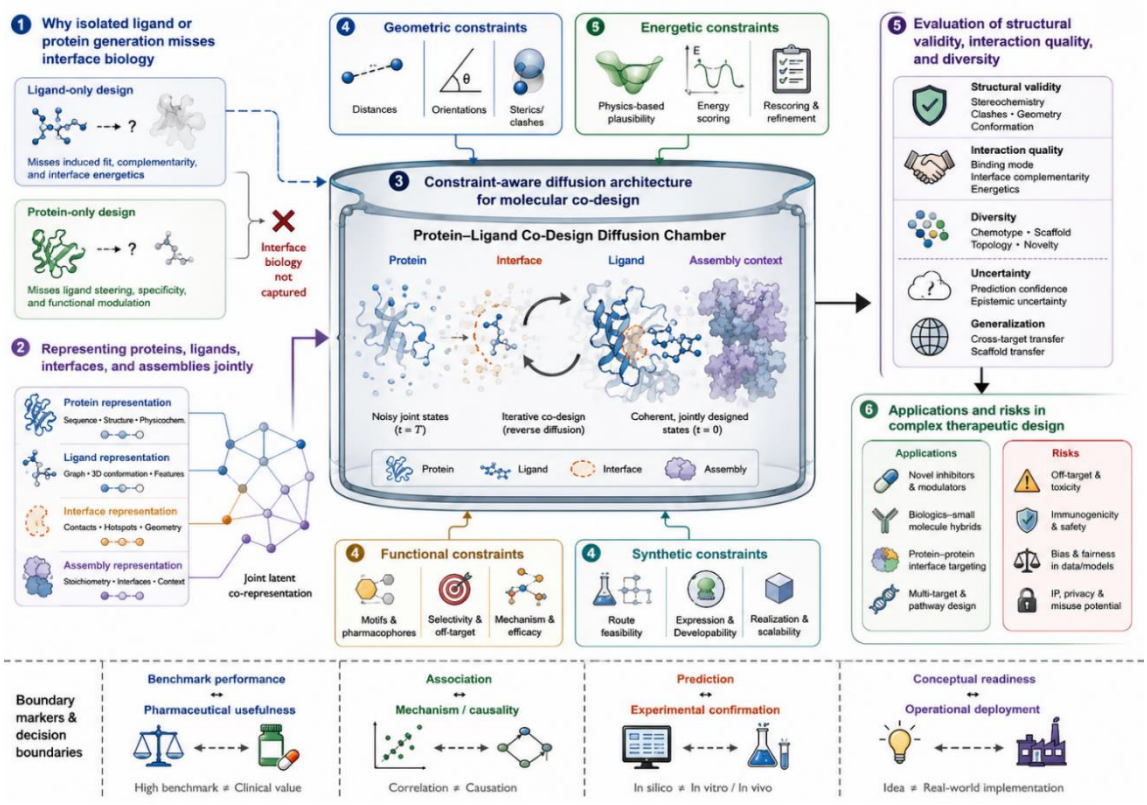

Figure 1. Protein–Ligand Co-Design Diffusion Chamber.

The figure is an original conceptual synthesis that organizes the article's central contribution across isolated ligand or protein generation misses interface biology, representing proteins, ligands, interfaces, and assemblies jointly, representing proteins, assemblies jointly, constraint-aware diffusion architecture for molecular co-design, geometric, energetic, functional, and synthetic constraints. Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, regulatory determination, clinical recommendation, or deployment-ready system.

Geometric, Energetic, Functional, and Synthetic Constraints

Geometric constraints define the immediately observable organization of a candidate system, including covalent geometry, chirality, bond lengths and angles, residue conformations, steric exclusion, intercomponent distances, directional contacts, motif placement, symmetry, and stoichiometry. These constraints are necessary but insufficient. Pose-validity analyses have shown that generated protein–ligand arrangements can appear convincing while containing implausible torsions, clashes, disconnected structures, or failures on unfamiliar protein sequences [17]. The proposed architecture therefore separates geometry used to guide generation from geometry used to reject physically incoherent outputs. Hard geometric constraints protect chemically mandatory relations, whereas uncertain interface preferences are treated as conditional or soft guidance. Passing geometric checks establishes only that the candidate is not disqualified by the assessed structural rules.

Energetic constraints address whether a proposed geometry is compatible with plausible molecular interactions and conformational states. Learned affinity models can extract recurring structural patterns, but their predictions do not necessarily generalize across targets, chemotypes, assay conditions, or unfamiliar interaction regimes [18]. For this reason, the architecture does not treat a docking score, neural affinity estimate, or approximate energy as a definitive objective. Instead, energetic assessment is distributed across independent channels that may include alternative conformations, physics-based rescoring, solvent-sensitive analysis, local strain inspection, and sensitivity to small structural perturbations. Disagreement among these channels should remain visible. A candidate with strong predicted affinity but severe geometric strain or unstable alternative poses should not be rescued by a favorable aggregate value.

Functional constraints specify what the molecular system is intended to do and which states must be promoted or avoided. Examples include preserving catalytic geometry, occluding an interaction surface, recruiting a ligase, stabilizing a desired conformational state, avoiding a homologous off-target, or generating multivalent engagement with a specified spatial arrangement. Functional constraints differ from structural constraints because the same apparent contact pattern can produce different biological consequences across environments, concentrations, cellular compartments, or kinetic regimes. The architecture therefore encodes function as a hypothesis linked to required structural relations and negative-design conditions, not as an automatically inferred property of a generated complex. Functional claims must advance only when supported by appropriate biochemical, biophysical, cellular, or system-level experiments.

Synthetic and realization constraints determine whether the proposed components can be produced in forms suitable for testing. Analysis of molecules generated by computational models has shown that formal validity and desirable predicted properties can coexist with poor or uncertain synthesizability [19]. Small-molecule outputs therefore require evaluation of precursor availability, transformation precedent, stereochemical control, protecting-group demands, route length, purification burden, and scale-dependent concerns. Route-planning systems can provide more informative evidence than a single accessibility score by linking target structures to plausible precursor and reaction sequences [20]. Protein components require a parallel realization analysis covering expression, folding, solubility, aggregation, proteolysis, modification, oligomerization, and assembly. These constraints can guide prioritization, but only laboratory synthesis, expression, purification, and characterization can establish realization.

Evaluation of Structural Validity, Interaction Quality, and Diversity

Evaluation must begin by rejecting the assumption that generative quality can be summarized by one universal score. Molecular-generation benchmarks distinguish validity, uniqueness, novelty, distributional similarity, diversity, and goal-directed optimization because each captures a different property and exposes different failure modes [21, 22]. A joint co-design system requires these dimensions but must extend them to proteins, interfaces, and assemblies. Evaluation should therefore record molecular graph validity, protein geometry, ligand pose integrity, interface complementarity, assembly topology, functional-constraint satisfaction, selectivity, synthetic feasibility, expression feasibility, and the diversity of complete interacting systems. Improvement in one dimension cannot compensate automatically for failure in another.

Uncertainty is a separate evaluation target rather than an optional confidence annotation. Comparative analyses of molecular uncertainty methods show that uncertainty estimates vary in calibration, scalability, and sensitivity to unfamiliar inputs [23]. The proposed architecture records uncertainty at several levels: molecular identity, conformation, interface geometry, assembly topology, constraint satisfaction, and downstream property prediction. It also distinguishes sampling variability from epistemic uncertainty and constraint uncertainty. A model may generate many similar candidates because the design region is genuinely narrow, because the sampler has collapsed, or because training data dominate the proposal. These explanations have different consequences and cannot be inferred from a confidence value alone.

Generalization must be assessed using splits and challenge sets that reflect the intended use. Ligand-based benchmarks can reward memorization when training and test examples share highly similar chemical or target patterns [24]. Equivalent risks arise in protein and interface design through sequence homology, structural-family overlap, recurring binding motifs, duplicated complexes, and retrospective use of structures that were unavailable at the relevant development time. Evaluation should therefore include scaffold-aware, target-family-aware, sequence-identity-aware, structure-cluster-aware, and time-aware partitions where appropriate. Nearest-neighbor disclosure and provenance analysis are necessary to determine whether apparent novelty or accuracy arises from interpolation near known examples.

The proposed Evidence, Validation, and Translation Boundary Observatory organizes these assessments as successive but nonautomatic evidentiary regions. Conceptual plausibility permits architectural reasoning; computational validation examines structural consistency, constraint satisfaction, diversity, uncertainty, and generalization; experimental confirmation tests whether components can be produced and behave as proposed; pharmaceutical interpretation asks whether the evidence is relevant to a consequential design decision. Advancement across these regions requires new evidence rather than reinterpretation of the same model output.

Applications and Risks in Complex Therapeutic Design

De novo protein binders provide a direct application for constrained joint design because their performance depends on both the designed protein and the target interface. Experimentally tested miniprotein inhibitors illustrate how computationally generated structures can progress to expression, binding assessment, structural analysis, and functional testing [25]. The proposed architecture may support such work by allowing epitope geometry, target conformation, binder backbone, sequence, negative-design partners, and expression constraints to be represented together. Its value would lie in generating better-structured hypotheses and exposing trade-offs before experimentation, not in replacing the experimental evidence needed to establish folding, affinity, specificity, mechanism, or efficacy.

Designed molecular assemblies expand the design object beyond a single interface. Modular protein components have been used to organize antibodies into nanocage-like architectures, demonstrating that designed interfaces, component geometry, and stoichiometric organization can be coordinated experimentally [26]. A joint diffusion architecture could represent these assembly variables explicitly and evaluate whether several local interfaces can coexist within the same global organization. Potential uses include multivalent binders, modular delivery systems, enzyme complexes, antibody assemblies, and programmable scaffolds. Major risks include incorrect stoichiometry, kinetic trapping, heterogeneous assembly, aggregation, payload interference, immunogenicity, and failure to reproduce the designed architecture under relevant formulation or biological conditions.

Targeted protein degradation is a particularly demanding test of molecular co-design because the target, degrader, linker, ligase, ternary interface, cellular context, exposure, and degradation machinery all contribute to the outcome. Reviews of PROTAC development emphasize that productive degradation cannot be reduced to binary affinity or isolated molecular properties [27]. The proposed architecture could organize ternary geometry, cooperative contacts, linker conformations, target and ligase states, and synthesis constraints within one generative scene. Nevertheless, predicted ternary stability would not

establish cellular degradation, selectivity, catalytic turnover, tissue exposure, resistance liability, or safety. The architecture should therefore be used, at most, to prioritize mechanistic hypotheses for orthogonal testing.

The same representational power that enables complex design also creates risks of persuasive but nonphysical outputs. Perturbational analysis of protein–ligand co-folding models indicates that models can produce confident-looking structures without consistently responding according to expected interaction physics [28]. Similar failure may occur when generative systems exploit training-set regularities, unrealistic protein flexibility, hidden data leakage, inadequate negative examples, or proxy objectives disconnected from the intended mechanism. Human review is required to challenge model assumptions, inspect constraint conflicts, select orthogonal validation methods, and decide whether evidence is sufficient for the proposed use. **Table 2** organizes the evidence, constructs, and interpretation boundaries needed to develop applications and risks in complex therapeutic design within the article's central argument.

Table 2. Evaluation, implementation, and research priorities arising from Co-Designing Proteins, Ligands, Binding Interfaces, and Molecular Assemblies through Constraint-Aware Generative Diffusion

Component or layer	Inputs	Core function	Expected output	Uncertainty or limitation	Validation requirement
Joint molecular scene	Protein, ligand, cofactor, interface, and assembly identities; coordinates; editable-state declarations	Preserve typed molecular entities and their cross-component relations in one generative state	Auditable representation of the interacting system	May omit solvent, dynamics, alternative states, or environmental variables	Representation-ablation studies and comparison with experimentally characterized systems
Protein state layer	Sequence, residue frames, backbone, side chains, modifications, conserved motifs	Represent fixed and editable protein variables	Protein structures and sequences consistent with declared constraints	Folding, expression, flexibility, and post-translational context may be incompletely modeled	Orthogonal structure prediction, expression, purification, stability, and structural experiments
Ligand state layer	Molecular graph, stereochemistry, conformers, protonation or tautomer alternatives, fragments	Generate or refine ligand identity and geometry	Chemically defined candidate ligand states	Chemical-state ambiguity and limited reaction-space coverage	Chemical validation, synthesis planning, laboratory synthesis, purity, and identity confirmation
Interface relation layer	Cross-component distances, orientations, contact types, surface descriptors	Make the binding interface an explicit design object	Candidate interaction geometry with typed relations	Static contacts may not represent solvent, kinetics, or conformational ensembles	Independent pose checks, mutational analysis, binding measurements, and structural confirmation
Assembly state layer	Component identity, stoichiometry, symmetry, topology, higher-order contacts	Coordinate multicomponent organization	Candidate assembly hypotheses	Pairwise-valid interfaces may be globally incompatible	Assembly-state characterization, stoichiometry measurements, microscopy, spectroscopy, or chromatography
Coupled equivariant denoiser	Noisy or partial joint scene; fixed masks; conditional structures	Update coordinates and selected identities without arbitrary frame dependence	Candidate interacting systems	Equivariance does not guarantee chemical or physical correctness	Controlled benchmarks, ablation studies, external targets, and prospective tests
Sequence–structure co-design head	Backbone geometry, interface context, sequence constraints	Couple residue identities to structural requirements	Candidate protein sequences for proposed structures	Sequence scoring may miss epistasis, folding kinetics, or expression liabilities	Expression, folding, stability, binding, and functional experiments
Constraint compiler	Human-defined requirements and computational descriptors	Translate requirements into hard constraints, guidance, exclusions, or tests	Machine-operable constraint specification	A proxy may poorly represent the underlying scientific requirement	Expert review, provenance reporting, and sensitivity analysis
Constraint-arbitration layer	Per-constraint feasibility, priorities, conflicts, and uncertainty	Expose incompatible requirements and prevent hidden compensation	Conflict records, Pareto alternatives, revise or reject decisions	Constraint priorities are application-dependent and may encode subjective choices	Predefined decision rules, expert adjudication, and reporting of unresolved conflicts

Geometric validation layer	Molecular and complex coordinates	Detect invalid local geometry, clashes, chirality errors, or assembly inconsistencies	Dimension-specific structural validity record	Passing structural checks does not establish binding or function	Independent geometric tools and experimental structures
Energetic assessment layer	Candidate structures, conformers, solvent assumptions, scoring methods	Evaluate physical plausibility without relying on one learned score	Comparative energetic evidence with disagreement retained	Estimates are method-, state-, and protocol-dependent	Orthogonal rescoring, simulation, perturbation studies, and experimental affinity data
Functional and negative-design layer	Desired mechanism, required states, off-targets, competing assemblies	Evaluate whether the candidate supports intended and excludes unwanted relations	Functional hypothesis and explicit failure conditions	Structural proxies may not capture cellular mechanism or selectivity	Biochemical, cellular, selectivity, and mechanism-focused assays
Synthetic feasibility layer	Molecular structure, reactions, precursors, route constraints	Assess whether ligand components have credible realization paths	Candidate retrosynthetic routes and accessibility evidence	Route existence does not establish yield, scale, safety, or manufacturing suitability	Medicinal-chemistry review and laboratory synthesis
Protein realization layer	Sequence, structure, physicochemical properties, oligomeric state	Assess expression, folding, solubility, aggregation, and assembly risks	Prioritized protein candidates with realization warnings	Computational filters cannot reproduce all host and formulation effects	Expression, purification, stability, aggregation, and assembly testing
Diversity and novelty layer	Generated candidate set and training-data provenance	Distinguish meaningful diversity from repetition or memorization	Structural, chemical, sequence, and assembly diversity profile	Novelty can be superficial or arise from database incompleteness	Similarity disclosure, clustering, and external-data comparison
Uncertainty and applicability layer	Predictions, ensembles, domain descriptors, constraint uncertainty	Separate confidence, calibration, and applicability limitations	Candidate-specific uncertainty record	Calibration may fail under target or distribution shift	Held-out calibration, external testing, prospective error analysis, and recalibration
Translation boundary layer	Full computational and experimental evidence package	Restrict claims to the evidence actually obtained	Bounded decision-use statement	Attractive early evidence may be overinterpreted as therapeutic readiness	Stage-appropriate developability, safety, pharmacology, and comparative studies
Provenance and audit layer	Data versions, structures, model versions, constraints, seeds, and reviewer actions	Enable reconstruction and accountability	Reproducible candidate and decision record	Proprietary or changing resources can obstruct independent audit	Versioned reporting, persistent identifiers, and independent reproduction

Limitations and Boundary Conditions

The principal conceptual limitation is that a joint representation does not make the underlying biological system fully observable. Static structures incompletely represent conformational ensembles, solvent reorganization, protonation coupling, kinetic barriers, induced fit, allostery, membrane context, macromolecular crowding, and cellular localization. A diffusion architecture may organize these uncertainties but cannot eliminate them. Generated geometry should therefore be interpreted as a state hypothesis rather than a complete thermodynamic or mechanistic description. Physical-validity checks can reject obvious structural failures, yet they cannot establish that a proposed state is populated under relevant conditions [17].

The evidence base also remains uneven across molecular modalities. Structural data favor experimentally tractable and historically studied proteins, ligands, and complexes; failed designs, inactive compounds, unsuccessful synthesis routes, expression failures, and incompatible assemblies are less consistently available. Learned energetic models may consequently reproduce dataset-specific regularities rather than transferable interaction principles [18]. Uncertainty methods can also appear calibrated within familiar distributions while failing on new targets or chemotypes [23]. The architecture therefore cannot establish universal applicability, and its validation would need to be repeated across molecular classes, target families, assembly types, and decision contexts.

Misuse could arise if an integrated architecture is interpreted as an integrated proof. Joint generation may produce visually coherent complexes that encourage stronger mechanistic or translational claims than the evidence permits. Benchmark leakage can exaggerate generalization [24], while perturbational failures can reveal that plausible outputs do not necessarily reflect learned physical understanding [28]. The proposed framework cannot establish binding affinity, causal mechanism, synthesis

success, expression, manufacturability, safety, clinical benefit, or regulatory acceptability. It can only organize candidate generation, expose constraint conflicts, and define the additional evidence required before a claim advances.

Research Agenda and Implementation Priorities

The first priority is to create evaluation resources that preserve the relational nature of the design problem. Datasets should connect molecular identities, structures, alternative conformations, assay conditions, affinities, functional outcomes, synthesis records, expression results, assembly states, and documented failures. Their splits should account for time, scaffold, target family, sequence identity, structural similarity, and repeated complexes. Synthetic information should include credible route evidence rather than only scalar accessibility labels [19], while route-planning methods should be evaluated on practical realization and not solely on reconstruction of known reactions [20]. Negative and failed examples are especially important for learning when constraints cannot be satisfied.

The second priority is prospective comparison of joint and component-wise design strategies. Blinded studies should ask whether simultaneous or iteratively coupled design produces better experimental hypotheses than workflows that generate one component while holding all others fixed. Evaluation should retain per-dimension outcomes rather than declaring a universal winner. For protein binders and assemblies, studies should connect generation to expression, folding, structural confirmation, affinity, selectivity, function, and assembly behavior, following the broader evidence progression demonstrated in experimentally tested protein and assembly design [25]. Comparable work is needed for multicomponent assemblies, where local interface quality must be evaluated against global organization [26].

The third priority is implementation governance. Every candidate should retain provenance for training data, input structures, molecular states, constraints, model versions, sampling settings, uncertainty estimates, validation tools, and human interventions. Constraint arbitration should report which requirements were satisfied, violated, relaxed, or left unresolved. Application-specific implementation should proceed in stages: retrospective reconstruction, leakage-controlled external evaluation, prospective computational challenge, design–make–test studies, and only then bounded decision support. Complex therapeutic applications such as targeted degradation require especially strict separation between structural hypotheses and cellular or translational evidence [27]. Perturbational challenge tests should remain part of validation because apparent structural coherence may conceal nonphysical model behavior [28].

Conclusion

Constraint-aware generative diffusion offers a conceptual route from isolated molecular generation toward co-design of interacting pharmaceutical systems. The proposed architecture makes proteins, ligands, interfaces, and assemblies jointly representable while retaining their distinct chemical, structural, and biological semantics. It also separates geometric, energetic, functional, synthetic, and realization constraints so that success in one dimension cannot silently compensate for failure in another. Its central contribution is therefore not a new claim of predictive accuracy, but an architectural definition of what joint molecular co-design should represent, how constraints and uncertainty should remain auditable, and how evidence should progress from structural plausibility to experimental and bounded pharmaceutical interpretation. The framework remains unimplemented as an integrated system and cannot establish mechanism, synthesizability, safety, therapeutic value, or operational readiness. Its usefulness will depend on leakage-resistant evaluation, explicit negative design, route and expression evidence, calibrated uncertainty, perturbational testing, and prospective design–make–test comparisons against simpler component-wise approaches.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Sánchez-Lengeling B, Aspuru-Guzik A. Inverse molecular design using machine learning: Generative models for matter engineering. Science. 2018;361(6400):360-5. doi:10.1126/science.aat2663
2. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17(2):97-113. doi:10.1038/nrd.2017.232
3. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-77. doi:10.1038/s41573-019-0024-5
4. 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(8016):493-500. doi:10.1038/s41586-024-07487-w
5. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein–protein interactions using AlphaFold2. Nat Commun. 2022;13(1):1265. doi:10.1038/s41467-022-28865-w

6. Gainza P, Sverrisson F, Monti F, Rodolà E, Boscaini D, Bronstein MM, et al. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. Nat Methods. 2020;17(2):184-92. doi:10.1038/s41592-019-0666-6
7. Gadd MS, Testa A, Lucas X, Chan KH, Chen W, Lamont DJ, et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. Nat Chem Biol. 2017;13(5):514-21. doi:10.1038/nchembio.2329
8. Zorba A, Nguyen C, Xu Y, Starr J, Borzilleri K, Smith J, et al. Delineating the role of cooperativity in the design of potent PROTACs for BTK. Proc Natl Acad Sci U S A. 2018;115(31). doi:10.1073/pnas.1803662115
9. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9. doi:10.1038/s41586-021-03819-2
10. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373(6557):871-6. doi:10.1126/science.abj8754
11. Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. Science. 2024;384(6693). doi:10.1126/science.adl2528
12. Humphreys IR, Pei J, Baek M, Krishnakumar A, Anishchenko I, Ovchinnikov S, et al. Computed structures of core eukaryotic protein complexes. Science. 2021;374(6573). doi:10.1126/science.abm4805
13. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. De novo design of protein structure and function with RFdiffusion. Nature. 2023;620(7976):1089-100. doi:10.1038/s41586-023-06415-8
14. Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, et al. Robust deep learning-based protein sequence design using ProteinMPNN. Science. 2022;378(6615):49-56. doi:10.1126/science.add2187
15. Schneuing A, Harris C, Du Y, Didi K, Jamasb AR, Igashov I, et al. Structure-based drug design with equivariant diffusion models. Nat Comput Sci. 2024;4(12):899-909. doi:10.1038/s43588-024-00737-x
16. Igashov I, Stärk H, Vignac C, Schneuing A, Garcia Satorras V, Frossard P, et al. Equivariant 3D-conditional diffusion model for molecular linker design. Nat Mach Intell. 2024;6(4):417-27. doi:10.1038/s42256-024-00815-9
17. Buttenschoen M, Morris GM, Deane CM. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. Chem Sci. 2024;15(9):3130-9. doi:10.1039/D3SC04185A
18. Volkov M, Turk JA, Drizard N, Martin N, Hoffmann B, Gaston-Mathé Y, et al. On the frustration to predict binding affinities from protein–ligand structures with deep neural networks. J Med Chem. 2022;65(11):7946-58. doi:10.1021/acs.jmedchem.2c00487
19. Gao W, Coley CW. The synthesizability of molecules proposed by generative models. J Chem Inf Model. 2020;60(12):5714-23. doi:10.1021/acs.jcim.0c00174
20. Segler MHS, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. Nature. 2018;555(7698):604-10. doi:10.1038/nature25978
21. Brown N, Fiscato M, Segler MHS, Vaucher AC. GuacaMol: Benchmarking models for de novo molecular design. J Chem Inf Model. 2019;59(3):1096-108. doi:10.1021/acs.jcim.8b00839
22. Polykovskiy D, Zhebrak A, Sanchez-Lengeling B, Golovanov S, Tatanov O, Belyaev S, et al. Molecular Sets (MOSES): A benchmarking platform for molecular generation models. Front Pharmacol. 2020;11:565644. doi:10.3389/fphar.2020.565644
23. Scalia G, Grambow CA, Pernici B, Li YP, Green WH. Evaluating scalable uncertainty estimation methods for deep learning-based molecular property prediction. J Chem Inf Model. 2020;60(6):2697-717. doi:10.1021/acs.jcim.9b00975
24. Wallach I, Heifets A. Most ligand-based classification benchmarks reward memorization rather than generalization. J Chem Inf Model. 2018;58(5):916-32. doi:10.1021/acs.jcim.7b00403
25. Cao L, Goreshnik I, Coventry B, Case JB, Miller L, Kozodoy L, et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. Science. 2020;370(6515):426-31. doi:10.1126/science.abd9909
26. Divine R, Dang HV, Ueda G, Fallas JA, Vulovic I, Sheffler W, et al. Designed proteins assemble antibodies into modular nanocages. Science. 2021;372(6537). doi:10.1126/science.abd9994
27. Békés M, Langley DR, Crews CM. PROTAC targeted protein degraders: The past is prologue. Nat Rev Drug Discov. 2022;21(3):181-200. doi:10.1038/s41573-021-00371-6
28. Masters MR, Mahmoud AH, Lill MA. Investigating whether deep learning models for co-folding learn the physics of protein–ligand interactions. Nat Commun. 2025;16(1):8854. doi:10.1038/s41467-025-63947-5