An AI-Driven Platform for Deconstructing and Engineering Biomolecular Recognition

Fodil Azzaz[1] & Jacques Fantini[1]

Aix-Marseille Univ, INSERM UA16, Marseille, France

Correspondence to: azzaz.fodil@gmail.com

Abstract

Understanding molecular interactions at the atomic level remains a central challenge across biology, medicine, and engineering. We introduce Perturbation Scanning (PS), an interrogative AI framework that actively deconstructs molecular interfaces. PS integrates graph-based representations of structures derived from molecular dynamics trajectories or Protein Data Bank files to systematically probe the electrostatic, hydrophobic, and steric contributions of each residue. To translate these insights into actionable design, we introduce the Intelligent Interface Optimization Scanner (IIOS), a standalone tool that generates energy-scored mutation proposals from interface maps. Together, PS and IIOS provide an integrated platform for dissecting and rationally engineering molecular interactions by resolving force-specific and stage-dependent contributions that are not directly accessible with existing computational approaches. Unlike traditional alanine scanning or free-energy methods such as MMPBSA—which provide only static or ensemble-averaged measures—PS delivers stage-resolved, force-wise decomposition of binding interfaces via a suite of targeted physicochemical perturbations applied to its AI model, directly quantifying each residue's mechanistic role.

Introduction

Understanding the physicochemical principles that govern molecular recognition requires moving beyond structural observation toward systematic interrogation. Molecular dynamics simulations capture interaction dynamics(1,2) , and machine learning models predict binding affinities, but these approaches remain fundamentally observational: they describe what happens during binding without actively testing which residues drive recognition through specific physical forces.

The central challenge is the transition from correlation to mechanism. Structural snapshots(3) and affinity measurements can identify residues associated with function, while free-energy calculations(4) provide ensemble-averaged contributions. However, these methods lack the resolution to decompose interfaces into their underlying electrostatic, hydrophobic, and steric components or to track how these contributions evolve across binding stages.

This limitation has broad practical consequences. In drug discovery(5), the inability to distinguish essential from peripheral contacts hinders rational optimization. In synthetic biology(6), engineering novel protein interfaces remains difficult without a clear understanding of the physical determinants of recognition. Even in well-studied systems such as neurotoxin–receptor interactions(7,8), differential binding can be observed, yet its physical basis often remains unresolved.

Current computational approaches—including molecular dynamics(1,2), free-energy calculations(4), and machine learning predictors— either remain observational or provide aggregated measures of binding. While these methods successfully describe interaction patterns or predict affinities, they lack a unified framework to resolve how distinct physicochemical forces contribute to binding across different temporal stages. Similarly, although protein design tools exist for stability or affinity optimization, they do not directly translate mechanistic interface profiles into ranked, force-specific mutation proposals. Here we close these gaps with Perturbation Scanning (PS) and the Intelligent Interface Optimization Scanner (IIOS). PS overcomes these limitations by integrating a graph neural network with a systematic perturbation engine, enabling residue-wise, force-specific, and stage resolved attribution of binding mechanisms.

PS employs a trained graph model as a surrogate system, enabling counterfactual interrogation via targeted physicochemical perturbations (e.g., charge reversal, hydrophobicity inversion). These perturbations independently modify electrostatic, hydrophobic, and steric properties across defined binding stages, directly quantifying residue and force-specific contributions. The framework operates on both molecular dynamics trajectories and static structures (PS-Lite), extending its utility across diverse systems. To convert mechanistic insight into actionable design, we integrate PS with the Intelligent Interface Optimization Scanner (IIOS), which transforms interface profiles into ranked, energetically scored mutation proposals. Together, PS and IIOS establish a unified paradigm for the mechanistic understanding and rational engineering of biomolecular interfaces (Fig. 1).

Results and Discussion

Rather than introducing new experimental datasets, we deliberately benchmark PS against multiple, independent, experimentally established binding phenomena across toxin–receptor and membrane systems, using agreement with known biological behavior as a stringent validation criterion.

The PS Framework: An AI Engine for Mechanistic Interrogation

We developed Perturbation Scanning (PS), a graph-based AI framework that transforms molecular interaction analysis from passive observation to active mechanistic interrogation. PS represents biomolecular systems as graphs in which atoms are encoded with physicochemical feature vectors—including electrostatics, hydrophobicity, mass, and structural context (9,10)—and edges represent covalent bonds or non-covalent contacts (<6 Å), identified via k-d tree algorithms (11). These graphs are processed by a graph convolutional network (GCN) (12) trained with frame interleaving, using batch normalization and dropout for regularization (13), to predict interface stability. The model learns hierarchical representations that integrate covalent and non-covalent interactions across conformational ensembles (**Fig. S1**).

Robustness was validated through 10 independent training trials with bootstrap confidence intervals (95% CI), assessing R^2, Spearman correlation, mean absolute error, and within-tolerance accuracy. The final model was selected based on balanced validation loss and test R^2, ensuring predictive reliability.

Configurable Temporal Staging Enables Stage-Resolved Mechanistic Deconstruction

A defining innovation of PS is its configurable temporal staging, which allows interaction analysis to be partitioned into user-defined binding stages. For clarity in this study, we divided trajectories into three equal-length regimes—early (initial binding events), mid (equilibrium interactions), and late (stabilized complex formation)—but the framework supports any number of stages, uneven time intervals, or unsupervised clustering to identify natural kinetic phases. This flexibility permits PS to model diverse dynamic processes, including stepwise complex assembly, induced-fit binding, and dissociation pathways, aligning the analysis with the underlying biochemical timeline.

Interface Perturbation Scanning Quantifies Mechanistic Contributions

To move beyond correlation toward mechanistic attribution, PS implements six perturbation strategies: electrostatic charge reversal, hydrophobic inversion, bulky side-chain substitution, π-stacking disruption, hydrogen-bond network breaking, and structural displacement. For each residue, the perturbation effect Δ is quantified as the absolute change in predicted interface stability: $\Delta = |P(\text{perturbed}) - P(\text{original})|$. Residues are ranked by the sum of Δ values across perturbations (Sum Δ), weighting both the magnitude and the frequency of interaction perturbation. This approach directly attributes binding contributions to specific physicochemical forces, answering not only *which* residues matter but *how* and *when* they contribute(14).

Computational Efficiency: Mechanistic Insight from a Single Trajectory

A key practical advantage of PS is its dramatic computational efficiency. Traditional methods for identifying critical residues typically require multiple independent MD simulations of mutant variants—each consuming hundreds of GPU hours and weeks of researcher time. For comprehensive analysis of interfaces like those studied here (20–30 residues across multiple binding stages), this translates to months of computational effort and substantial financial cost. In contrast, PS extracts comparable mechanistic insights relevant to residue-level and force-specific contributions through *in silico* perturbations applied to a single wild-type trajectory, avoiding the need for separate mutant simulations and reducing computational requirements by approximately N-fold (where N is the number of residues studied). This efficiency gain enables rapid comparative analysis across multiple systems—as demonstrated below—that would be computationally prohibitive through brute-force mutation simulations, democratizing detailed mechanistic analysis for laboratories without access to supercomputing resources.

BoNT/B1–SYT1 and BoNT/B1–SYT2 Systems Reveal Distinct Assembly Architectures

Applied to the BoNT/B1–SYT1 and BoNT/B1–SYT2 neurotoxin–receptor systems, PS captured fundamentally different dynamic landscapes. For the flexible SYT1 complex, the model achieved a mean test R^2 of 0.81 ± 0.035 with Spearman $\rho = 0.90 \pm 0.013$, reflecting broad conformational fluctuations. In contrast, the rigid SYT2 complex yielded $R^2 \approx 0$ due to minimal variance across the trajectory, yet maintained substantial rank correlation ($\rho = 0.704 \pm 0.031$), demonstrating PS's ability to adapt to systems with varying degrees of conformational freedom (Fig. S2A–L).

Applying the PS perturbation suite revealed that SYT1 employs a GT1b-dependent compensatory assembly strategy (Fig. 4A–B). Initial SYT1–GT1b interactions weaken over time ($\Delta = 0.472 \rightarrow 0.144$), compensated by strengthening BoNT/B1–GT1b contacts ($\Delta = 0.647 \rightarrow 1.123$), indicating a dynamic handoff mechanism where the ganglioside transitions from SYT1 recruitment to BoNT/B1 stabilization (Fig. 2C). This GT1b-mediated handoff mechanistically explains the experimentally established requirement for ganglioside presence in BoNT/B1 binding to SYT1, revealing how this dependence operates dynamically during assembly (15). Key residues driving this compensation include LYS-1187/1188 and GLU-1189 at the BoNT/B1-SYT1 interface, and R1242 engaging GT1b's sialic acid and N-acetylgalactosamine moieties (Fig. 2D).

In contrast, SYT2 exhibits a modular, autonomous binding architecture, with a stronger BoNT/B1 interface (Sum $\Delta = 4.020$) and a recoverable U-shaped SYT2–GT1b engagement. This autonomy is stabilized by K1187/1188, K46, and N56, with ganglioside contacts maintained through a hydrophobic patch (A64, L65, I68) (Fig. 2E).

Together, these findings illustrate how PS decodes not only interaction patterns but the physical mechanisms underlying differential receptor specificity. PS provides a mechanistic deconstruction of known experimental observations: SYT1's ganglioside dependence emerges from a GT1b-bridged compensatory assembly, whereas SYT2 engages via a direct, protein-centric modular interface.

BoNT/A1–SV2A and BoNT/A1–SV2C Systems Unveil Divergent Temporal Engagement Strategies

Expanding to BoNT/A1–SV2A and BoNT/A1–SV2C complexes, PS uncovered strikingly different binding dynamics. The SV2C interface demonstrated progressive stabilization, with perturbation sensitivity increasing from Δ = 2.962 (early) to 3.368 (late), indicating continuous optimization. Conversely, the SV2A complex exhibited transient instability, starting strong (Δ = 3.260 early), dropping sharply at mid-stage (2.508), then partially recovering (2.714 late), suggesting rearrangement or competitive displacement before weaker final contacts form (Fig. 3A).

Residue-level analysis highlighted Thr-1145 as a dominant contributor in SV2C (Δ = 0.4938) versus a much weaker role in SV2A (0.1416), with Met-1143 and Arg-1293 also showing enhanced contributions in the high-affinity complex (Fig. 3B). Interaction-type profiling revealed that SV2C binding is hydrogen-bond-dominated (45% of interface strength), while SV2A exhibits a more balanced distribution (35% H-bonds, 30% hydrophobic, 20% electrostatic) (Fig. 3C). These findings align with and mechanistically explain the experimentally established higher binding affinity of BoNT/A1 for SV2C over SV2A(16). PS reveals the biophysical mechanism behind SV2C's superior affinity: through a progressively optimized hydrogen-bond network that strengthens across binding stages, while SV2A relies on transient, balanced interactions that fail to stabilize over time.

Plasmolipin–Lipid Interactions: Lipid-Specific Resolution in Membrane Environments

Beyond protein–protein systems PS provides per-lipid resolution in analyzing membrane protein environments, enabling quantitative comparison of individual lipid contributions. Applied to plasmolipin embedded in a lipid-raft membrane (POPC/cholesterol/GM1) versus a pure POPC bilayer(17), PS revealed ~4-fold higher interaction strengths in the raft system (early Δ = 33.99 vs. 8.37 in POPC) (Table S1). Strikingly, ceramide (CER160), despite constituting only 13.1% of contact frequency, contributed 35.1% of the total Δ score (16.48 Δ), demonstrating a disproportionate stabilizing role. In contrast, POPC (44.3% contact frequency) accounted for only 21.4% (10.06 Δ). Ganglioside polar components and cholesterol collectively contributed 39.3% of the total interaction score (Table S2). This lipid-resolved insight moves beyond phenomenological descriptions of "raft stabilization" to quantify how specific lipid chemistries orchestrate membrane protein stability

through stage-dependent interactions, establishing a quantitative basis for understanding lipid selectivity and therapeutic targeting in membrane biology.

PS-Lite: Accessible Interface Analysis for Static Structures

To broaden accessibility, we developed PS-Lite, a deterministic, non–machine-learning variant designed for rapid and interpretable analysis of static structures. PS-Lite applies the same physicochemical feature engineering as full PS but replaces the neural network with rule-based scoring based on contact density and residue properties. Validation on the Barnase–Barstar complex (PDB: 1BRS) correctly identified experimentally critical residues (Barnase Arg-59, Arg-87; Barstar Asp-35, Asp-39). Applied to BoNT/A1–SV2C (PDB: 6ES1), it highlighted key interface residues including SV2C Lys-558, Phe-563, and Asn-559, and BoNT/A1 Tyr-1122, Tyr-1149, and Thr-1145/1146. While lacking temporal resolution, PS-Lite provides rapid, interpretable interface mapping suitable for mutagenesis guidance, preliminary screening, and educational use, extending the framework's utility to researchers without MD capabilities.

IIOS Identifies Actionable Mutation Opportunities

To translate PS-derived interface insights into rational design, we applied the Intelligent Interface Optimization Scanner (IIOS). For the Barnase–Barstar complex, IIOS generated 48 prioritized mutation suggestions, including SER38→ARG (ΔΔG_pred = −2.5 kcal/mol) and ARG59→PHE (ΔΔG_pred = −2.3 kcal/mol) to establish new salt bridges and π-stacking interactions. Proximity scanning further identified under-utilized interfacial positions such as LEU89→ARG and THR100→PHE, highlighting the potential to create novel contacts (Fig. 4A).

Applied to a BoNT/B1–SYT2 molecular dynamics trajectory, IIOS proposed 34 mutations, with ASN1175→ARG and GLN1176→ARG predicted to strengthen electrostatic complementarity near SYT2 Glu-rich regions. Aromatic enhancements (e.g., THR1182→PHE) were suggested to reinforce π-stacking with SYT2 Phe-48, while proximity scanning revealed opportunities for new π-stacking interactions at near-interface residues such as ILE1259→PHE/TYR (Fig. 4B). These proposals illustrate how IIOS systematically converts structural and dynamic interface profiles into energetically scored design blueprints, bridging mechanistic understanding with practical protein engineering.

PS as a Complementary Analytical Layer for Mechanistic Interface Analysis

The development of PS represents a shift from observational to interrogative computational structural biology. Unlike molecular dynamics simulations—which describe but do not explain—or free-energy methods that provide ensemble-averaged contributions, PS actively quantifies the role of individual residues through systematic *in silico* perturbations. Unlike machine-learning predictors that correlate structure with affinity, PS attributes binding contributions to specific electrostatic, hydrophobic, and steric forces across user-defined kinetic

stages. This stage-resolved, force-wise deconstruction has no direct counterpart in current computational toolkits. By integrating graph-based representation, strategic frame interleaving, and multi-faceted *in silico* mutagenesis, PS moves beyond identifying important residues to deconstructing their stage-wise physicochemical contributions. Benchmarking across four neurotoxin–receptor systems spanning two toxin families (BoNT/B1 with α-helix recognition and BoNT/A1 with β-strand engagement) demonstrates PS's generality and robustness. The framework successfully captured compensatory assembly (SYT1), modular autonomy (SYT2), progressive stabilization (SV2C), and transient instability (SV2A), revealing that molecular recognition is not merely a static complementarity but an orchestrated sequence of physical interactions.

Biological Implications and Therapeutic Design

The differential binding strategies decoded by PS have direct implications for understanding botulism pathogenesis and developing targeted interventions. For example, SYT1's ganglioside-dependent compensatory mechanism suggests that engineered soluble decoys could disrupt the handoff step, while SYT2's protein-centric modular interface might be blocked by high-affinity peptide inhibitors. Similarly, SV2C's progressive hydrogen-bond network provides a template for designing stable binders, whereas SV2A's transient interface may require allosteric disruptors. More broadly, PS reveals that binding sequence—the temporal order of interaction formation—may be as important as binding affinity for biological specificity, aligning with emerging systems-biology perspectives on stage-coordinated function. PS does not claim experimental causality; rather, it provides a controlled counterfactual interrogation of physicochemical contributions that can guide and rationalize experimental design.

Limitations and Future Directions

While PS represents a significant advance, several limitations warrant consideration. First, perturbations are computational; experimental validation of key predictions across multiple systems is needed. Second, analysis of individual trajectories could be strengthened by integrating multiple replicas for more robust statistics. Third, stage definition, though flexible, requires user input; automated clustering approaches could further streamline analysis. Fourth, future applications to systems with large conformational changes, disordered regions, or multi-domain rearrangements will further test the framework's robustness. Directions for extension include: (1) incorporating experimental data via multi-modal learning, (2) developing adaptive stage-segmentation algorithms, (3) extending perturbation strategies to post-translational modifications and allosteric effects, and (4) establishing standardized benchmarks for stage-resolved interaction analysis across broader molecular classes. The modular architecture of PS ensures that these extensions can be implemented without fundamental redesign. PS is intended as a hypothesis-

generating and mechanistic interpretation framework, not a replacement for experimental validation.

Conclusion: From Observational Science to Engineering Discipline

Perturbation Scanning (PS) provides a generalizable, computationally efficient framework for deconstructing biomolecular interfaces into their fundamental physical components across user-definable binding stages. By revealing not only *which* residues participate but *how* and *when* they contribute through specific forces, PS transforms our capacity to decode the complex language of molecular recognition. In doing so, PS moves beyond the aggregated statistics of free-energy calculations and the static approximations of traditional mutagenesis tools, introducing a stage-aware analytical framework for mechanistic biomolecular engineering. Coupled with the standalone design tool IIOS, the pipeline closes the loop from mechanistic insight to rational interface engineering. As structural biology increasingly captures transient states and AI generates dynamic ensembles, frameworks like PS will be essential for extracting mechanistic understanding from complex datasets. By making stage-dependent mechanisms interpretable and computationally accessible, PS shifts molecular recognition from an observational science to an engineering discipline—enabling researchers not only to understand but to actively orchestrate biomolecular interactions across diverse biological contexts.

Author Contributions.

F.A.: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

J.F.: Supervision, Writing – Original Draft, Writing – Review & Editing, Resources.

Competing Interests.

F.A. is the creator and copyright holder of the PS software suite, which is made available under a custom open-source license. J.F. declares no competing interests.

Data and Materials Availability:

The Perturbation Scanning (PS) software suite, including PS-Lite and IIOS, is publicly available as well-documented, user-ready source code at https://github.com/fodil13.

The repository includes tutorials, example configurations, and centralized parameter files, enabling researchers to apply the framework to their own systems with minimal setup. The code is released under a custom open-source license that permits free academic use; commercial use requires a separate license. All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials.

References

1. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from

laptops to supercomputers. SoftwareX [Internet]. 2015;1–2:19–25. Available from: https://www.sciencedirect.com/science/article/pii/S2352711015000059

2. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem [Internet]. 2005 Dec 1;26(16):1781–802. Available from: https://doi.org/10.1002/jcc.20289

3. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res [Internet]. 2000 Jan 1;28(1):235–42. Available from: https://doi.org/10.1093/nar/28.1.235

4. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K. Free Energy Calculation from Steered Molecular Dynamics Simulations Using Jarzynski's Equality.

5. Congreve M, Carr R, Murray C, Jhoti H. A 'Rule of Three' for fragment-based lead discovery? Drug Discov Today [Internet]. 2003;8(19):876–7. Available from: https://www.sciencedirect.com/science/article/pii/S1359644603028319

6. Benner SA, Sismour AM. Synthetic biology. Nat Rev Genet [Internet]. 2005;6(7):533–43. Available from: https://doi.org/10.1038/nrg1637

7. Azzaz F, Hilaire D, Fantini J. Structural basis of botulinum neurotoxin serotype A1 binding to human SV2A or SV2C receptors. Chem Biol Interact [Internet]. 2023;373:110384. Available from: https://www.sciencedirect.com/science/article/pii/S0009279723000510

8. Azzaz F, Hilaire D, Fantini J. Structural Basis of Botulinum Toxin Type F Binding to Glycosylated Human SV2A: In Silico Studies at the Periphery of a Lipid Raft. Biomolecules [Internet]. 2022;12(12). Available from: https://www.mdpi.com/2218-273X/12/12/1821

9. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods [Internet]. 2017;14(1):71–3. Available from: https://doi.org/10.1038/nmeth.4067

10. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol [Internet]. 1982;157(1):105–32. Available from: https://www.sciencedirect.com/science/article/pii/0022283682905150

11. Maneewongvatana S, Mount DM. Analysis of Approximate Nearest Neighbor Searching with Clustered Point Sets ⋆. 1999.

12. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural Message Passing for Quantum Chemistry. 2017 Jun 12; Available from: http://arxiv.org/abs/1704.01212

13. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015 Mar 2; Available from: http://arxiv.org/abs/1502.03167

14. Fantini J, Matveeva M, Lefebvre M, Chahinian H. What Is life? Rethinking Biology in Light of Fundamental Parameters. Life [Internet]. 2024;14(3). Available from: https://www.mdpi.com/2075-1729/14/3/280

15. Azzaz F, Mazzarino M, Chahinian H, Yahi N, Scala C Di, Fantini J. Structure of the Myelin Sheath Proteolipid Plasmolipin (PLLP) in a Ganglioside-Containing Lipid Raft. Front Biosci (Landmark Ed). 2023;28(8):157.

Figure legends

Fig 1. Overview of the Perturbation Scanning (PS) and Intelligent Interface Optimization Scanner (IIOS) framework. Left, a biomolecular interface is interrogated beyond structural observation, highlighting the central question of how specific residues contribute to recognition. Center, Perturbation Scanning (PS) deconstructs molecular interfaces by applying systematic, force-specific in silico perturbations to a graph-based AI surrogate model trained on molecular dynamics trajectories or static structures. Residue-wise perturbations (e.g., electrostatic, hydrophobic, steric) are applied across user-defined binding stages, enabling mechanistic attribution of how and when individual residues contribute to interface stability. Bottom center, Intelligent Interface Optimization Scanner (IIOS) translates PS-derived physicochemical interface maps into actionable design proposals by scanning interfacial and near-interface residues and suggesting energetically scored mutations that improve physical–chemical complementarity. Right, representative application areas include rational antibody and protein therapeutic design, allosteric and orthosteric drug discovery, enzyme engineering for biocatalysis and biosynthesis, and membrane protein engineering and synthetic biology. Together, PS and IIOS form an integrated platform that closes the loop from mechanistic interface deconstruction to rational biomolecular engineering.

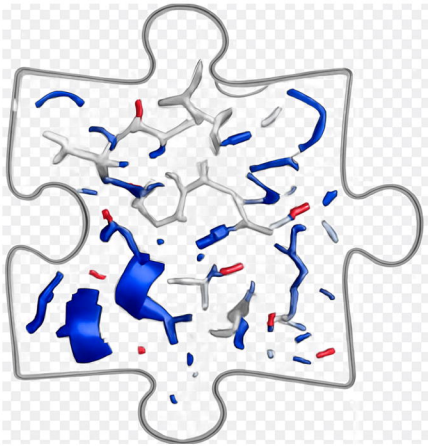
Fig 2. Comparative Interface for BoNT/B1-SYT1-GT1b and BoNT/B1-SYT2-GT1b systems. Total interface strength comparison showing SYT2's stronger autonomous protein-protein binding (blue) versus SYT1's compensatory GT1b bridging (orange) (A). Temporal dynamics reveal SYT1's compensatory handoff (dashed lines) versus SYT2's modular stability (solid lines) (B). Network architecture models illustrate compensatory (SYT1) versus modular (SYT2) assembly strategies; edge thickness ∝ interface strength (C). Top residue contributors for SYT1 and SYT2 systems, ranked by Sum Δ and colored by perturbation type sensitivity (D-E).

Fig 3. Perturbation Scanning Analysis of BoNT/A1 Binding to SV2A vs SV2C. Bar plots show the Sum Δ (interface perturbation sensitivity) for SV2C (blue) and SV2A (purple) complexes across early, mid and late binding stages (A). Horizontal bar chart siplaying the mean Sum Δ values for the top five contributing residues from

BoNT/A1 in each complex (B). Pie charts comparing the percentage contribution of different physicochemical interaction types to total interface strength (C). Hsd is for histidine, BGALNA is for β-D-N-acetylgalactosamine, and BGAL is for β-galactose. The numbers reflect their position in the sequence (from 1 to 7).

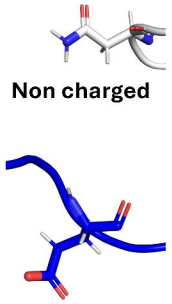
Fig 4: Exemple of raw output provided by Intelligent Interface Optimization Scanner (IIOS) for Barnase-Barstar (A) and BoNT/B1-SYT2 complex (B).

Perturbation Scanning (PS)

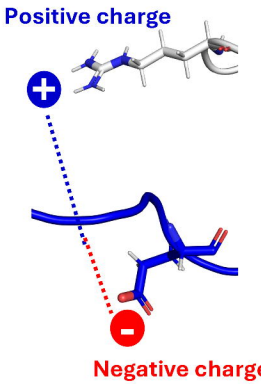


Disect to understand

Intelligent Interface Optimisation Scanning (IIOS)



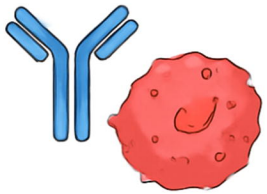
Before IIOS:
No Physical-Chemical Complementarity



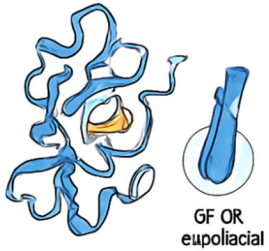
After IIOS:
Propose a Physical-Chemical Complementarity

APPLICATION Areas

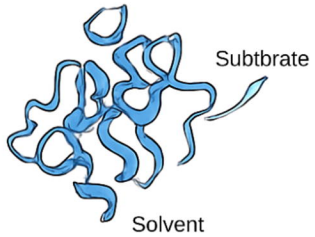
Rational Antibody & Protein Therapeutics Design



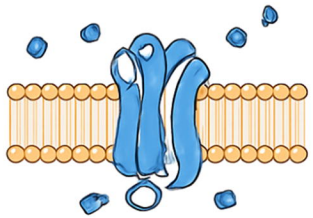
Allosteric & Orthosteric Drug Discovery

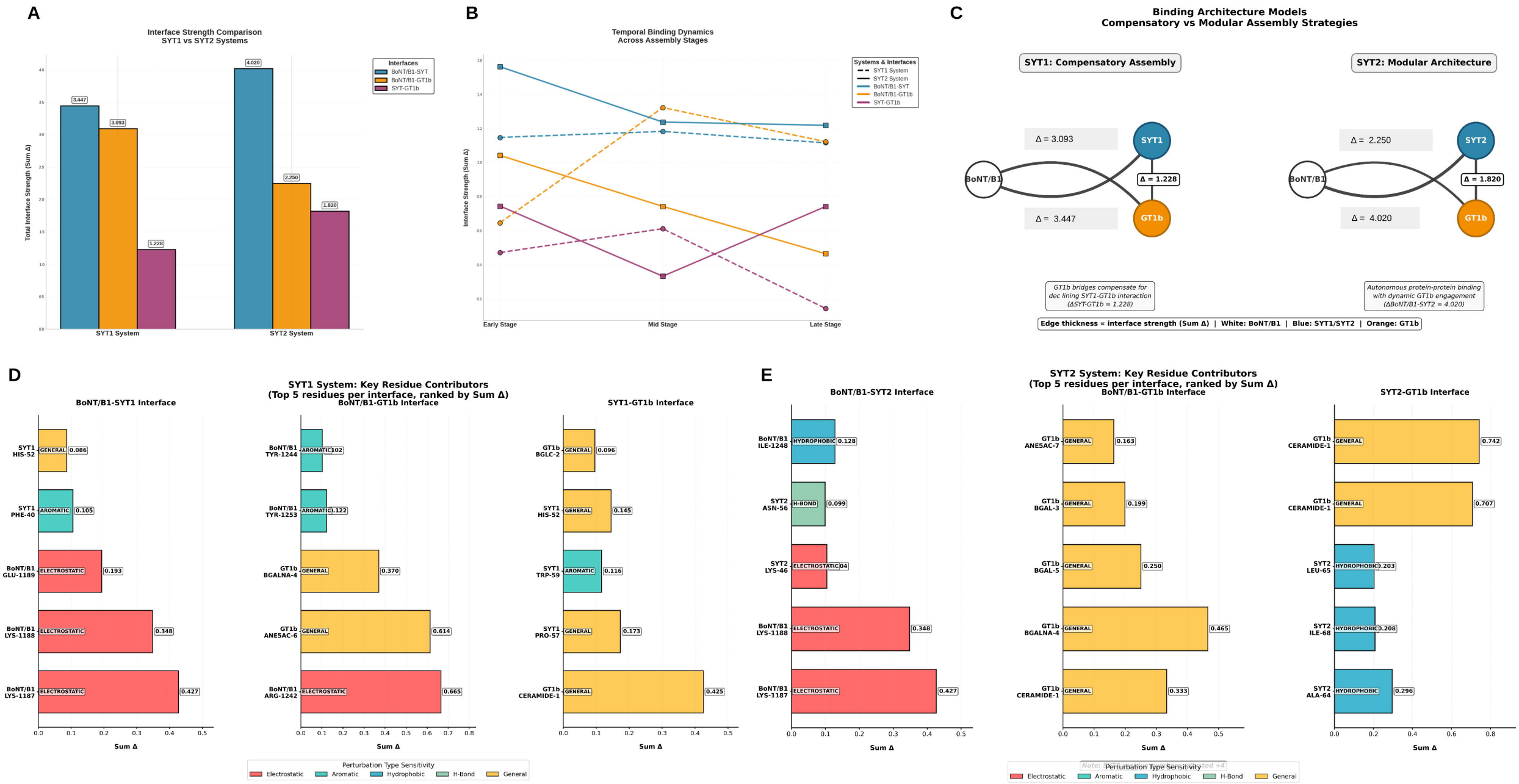


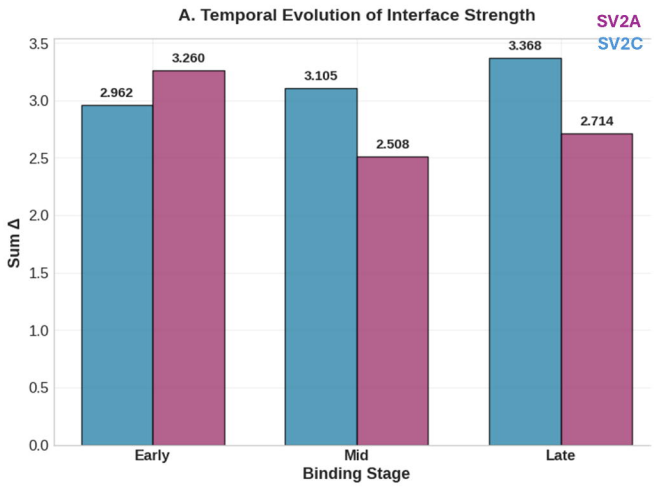
Enzyme Engineering for Biocatalysis & Biosynthesis

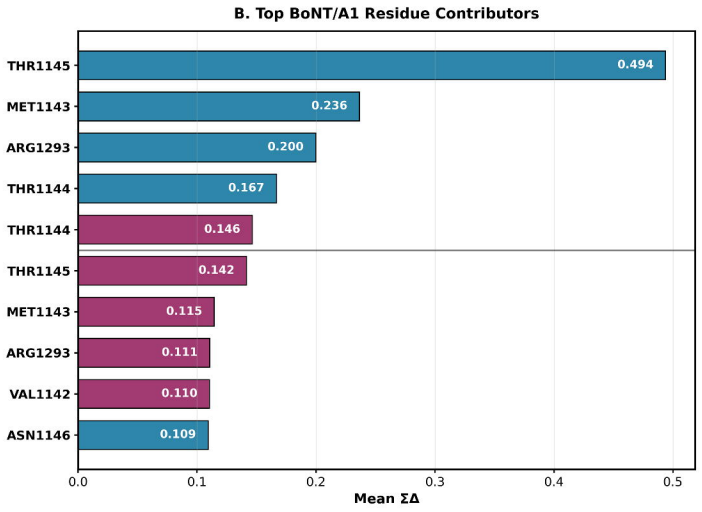


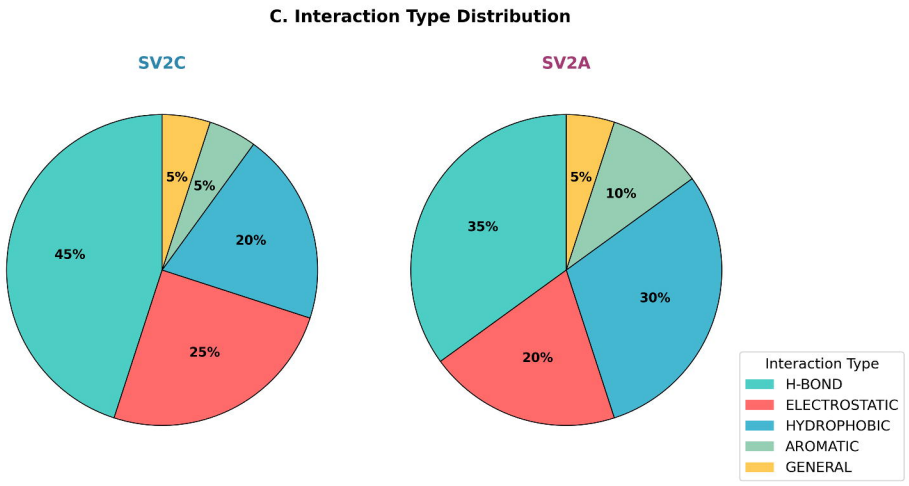
Membrane Protein Engineering & Synthic Biology




A
Interface Strength Comparison
SYT1 vs SYT2 Systems
Total Interface Strength (Sum Δ)
3.447
3.093
1.228
4.020
2.250
1.820
SYT1 System
SYT2 System
Interfaces
BoNT/B1-SYT
BoNT/B1-GT1b
SYT-GT1b
B
Temporal Binding Dynamics
Across Assembly Stages
Interface Strength (Sum Δ)
Early Stage
Mid Stage
Late Stage
Systems & Interfaces
SYT1 System
SYT2 System
BoNT/B1-SYT
BoNT/B1-GT1b
SYT-GT1b
C
Binding Architecture Models
Compensatory vs Modular Assembly Strategies
SYT1: Compensatory Assembly
SYT2: Modular Architecture
Δ = 3.093
SYT1
BoNT/B1
Δ = 1.228
Δ = 3.447
GT1b
Δ = 2.250
SYT2
BoNT/B1
Δ = 1.820
Δ = 4.020
GT1b
GT1b bridges compensate for dec lining SYT1-GT1b interaction (ΔSYT-GT1b = 1.228)
Autonomous protein-protein binding with dynamic GT1b engagement (ΔBoNT/B1-SYT2 = 4.020)
Edge thickness ∝ interface strength (Sum Δ) | White: BoNT/B1 | Blue: SYT1/SYT2 | Orange: GT1b
D
SYT1 System: Key Residue Contributors
(Top 5 residues per interface, ranked by Sum Δ)
BoNT/B1-SYT1 Interface
SYT1 HIS-52 GENERAL 0.086
SYT1 PHE-40 AROMATIC 0.105
BoNT/B1 GLU-1189 ELECTROSTATIC 0.193
BoNT/B1 LYS-1188 ELECTROSTATIC 0.348
BoNT/B1 LYS-1187 ELECTROSTATIC 0.427
BoNT/B1-GT1b Interface
BoNT/B1 TYR-1244 AROMATIC 0.102
BoNT/B1 TYR-1253 AROMATIC 0.122
GT1b BGALNA-4 GENERAL 0.370
GT1b ANE5AC-6 GENERAL 0.614
BoNT/B1 ARG-1242 ELECTROSTATIC 0.665
SYT1-GT1b Interface
GT1b BGLC-2 GENERAL 0.096
SYT1 HIS-52 GENERAL 0.145
SYT1 TRP-59 AROMATIC 0.116
SYT1 PRO-57 GENERAL 0.173
GT1b CERAMIDE-1 GENERAL 0.425
Sum Δ
Perturbation Type Sensitivity
Electrostatic
Aromatic
Hydrophobic
H-Bond
General
E
SYT2 System: Key Residue Contributors
(Top 5 residues per interface, ranked by Sum Δ)
BoNT/B1-SYT2 Interface
BoNT/B1 ILE-1248 HYDROPHOBIC 0.128
SYT2 ASN-56 H-BOND 0.099
SYT2 LYS-46 ELECTROSTATIC 0.104
BoNT/B1 LYS-1188 ELECTROSTATIC 0.348
BoNT/B1 LYS-1187 ELECTROSTATIC 0.427
BoNT/B1-GT1b Interface
GT1b ANE5AC-7 GENERAL 0.163
GT1b BGAL-3 GENERAL 0.199
GT1b BGAL-5 GENERAL 0.250
GT1b BGALNA-4 GENERAL 0.465
GT1b CERAMIDE-1 GENERAL 0.333
SYT2-GT1b Interface
GT1b CERAMIDE-1 GENERAL 0.742
GT1b CERAMIDE-1 GENERAL 0.707
SYT2 LEU-65 HYDROPHOBIC 0.203
SYT2 ILE-68 HYDROPHOBIC 0.208
SYT2 ALA-64 HYDROPHOBIC 0.296
Sum Δ
Perturbation Type Sensitivity
Electrostatic
Aromatic
Hydrophobic
H-Bond
General


A. Temporal Evolution of Interface Strength
SV2A
SV2C
Sum Δ
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2.962
3.260
3.105
2.508
3.368
2.714
Early
Mid
Late
Binding Stage


B. Top BoNT/A1 Residue Contributors
THR1145
MET1143
ARG1293
THR1144
THR1144
THR1145
MET1143
ARG1293
VAL1142
ASN1146
0.494
0.236
0.200
0.167
0.146
0.142
0.115
0.111
0.110
0.109
0.0
0.1
0.2
0.3
0.4
0.5
Mean ΣΔ


C. Interaction Type Distribution
SV2C
45%
25%
20%
5%
5%
SV2A
35%
20%
30%
10%
5%
Interaction Type
H-BOND
ELECTROSTATIC
HYDROPHOBIC
AROMATIC
GENERAL

A

🔵 RANK #1 [CURRENT PARTNER]

📍 TARGET: PROA-SER-38
- Mutation: SER→ARG
- Expected ΔΔG: -2.50 kcal/mol
- Pairing: 100.0%
- Interaction: Salt Bridge (avg distance: 2.8Å)
- Confidence: 87%

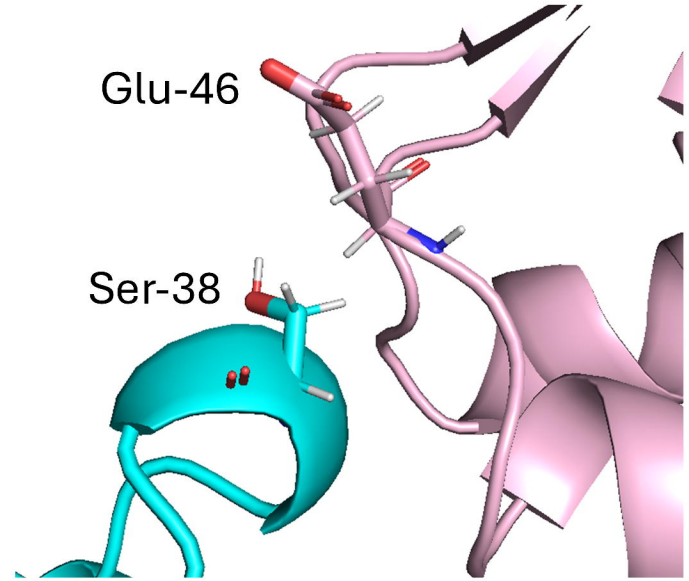
} Description of the mutation and predicted effect

REASON:
Introduce salt bridge with GLU (charge: -1) at 2.8Å } Explanation of « why this mutation »

🤝 PARTNER(S):
1. PROD-GLU-46 - GLU
   Distance: 2.8Å
   Type: salt bridge

} Information about the partner

PRIORITY: HIGH - Strong improvement expected



B

🔵 RANK #1 [CURRENT PARTNER]

📍 TARGET: PROA-ASN-1175
- Mutation: ASN→ARG
- Expected ΔΔG: -2.50 kcal/mol
- Pairing: 55.1%
- Interaction: Salt Bridge (avg distance: 3.0Å)
- Confidence: 86%

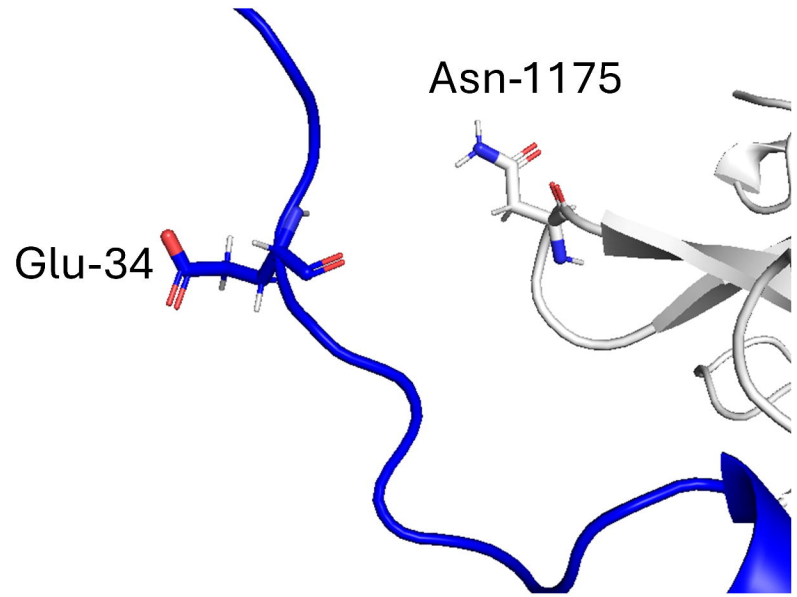
} Description of the mutation and predicted effect

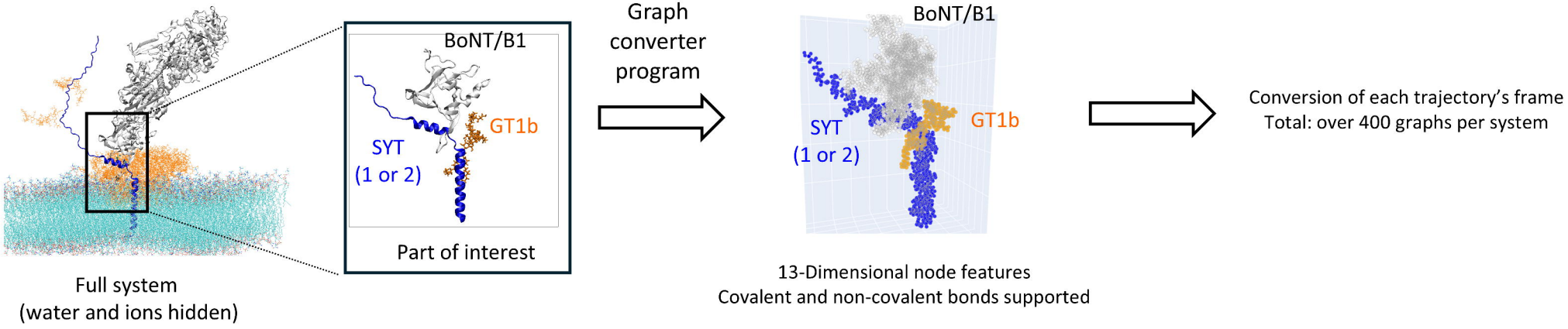
💡 REASON:
Introduce salt bridge with GLU (charge: -1) at 3.0Å } Explanation of « why this mutation »

PARTNER(S):
1. PROB-GLU-34 - GLU
   Distance: 3.0Å
   Type: salt bridge

} Information about the partner

PRIORITY: HIGH - Strong improvement expected





Architecture

INPUT: [13D Features]

↓

GCN LAYERS: 128 → 64 → 32

(BatchNorm + ReLU + Dropout)

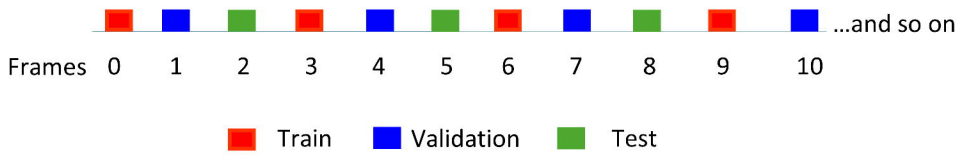
↓

INTERFACE FOCUS: BoNT/B1-SYT and BoNT/B1-GT1b and SYT-GT1b

↓

OUTPUT: Trained Model for Comparative Binding Mechanism Analysis

Frame Interleaving Strategy



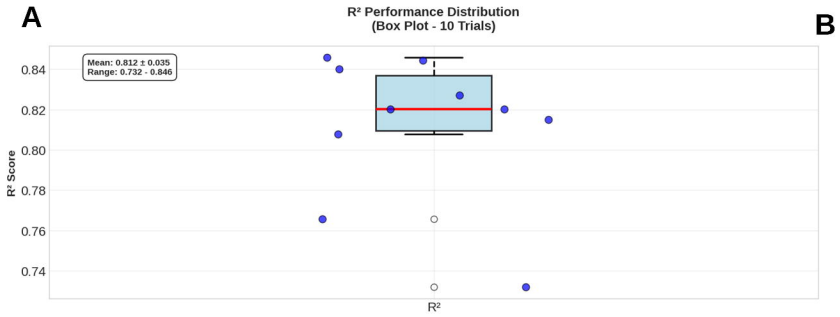
Multi-run validation

Trial 1: Different Random Seed $R^2 = 0.XX \pm 0.YY$ Spearman = 0.XX
Trial 2: Different Random Seed $R^2 = 0.XX \pm 0.YY$ Spearman = 0.XX
Trial 3: Different Random Seed $R^2 = 0.XX \pm 0.YY$ Spearman = 0.XX
...
Trial 10: Different Random Seed $R^2 = 0.XX \pm 0.YY$ Spearman = 0.XX

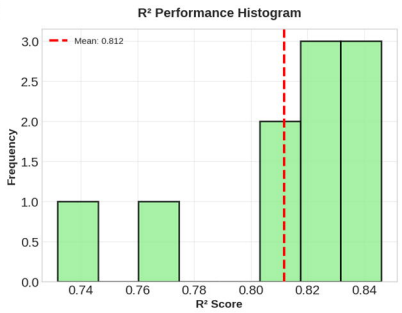
FINAL: $R^2 = 0.XX \pm 0.YY$ (95% CI) - Pure Model Variance

BoNT/B1-SYT1 System

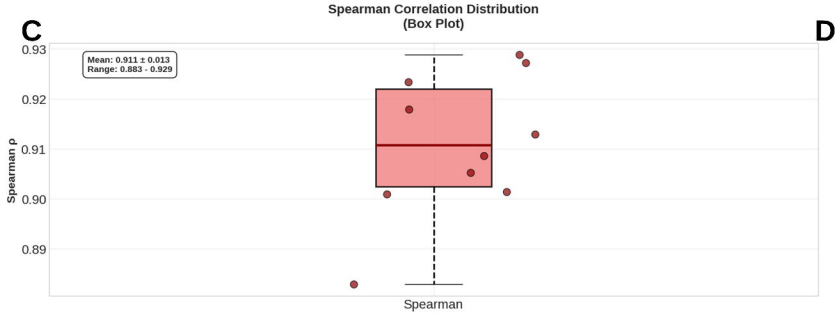
A



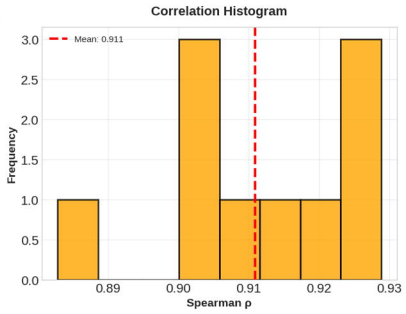
B



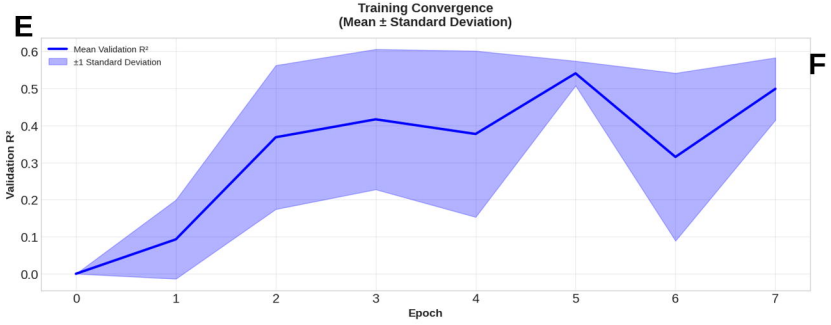
C



D



E



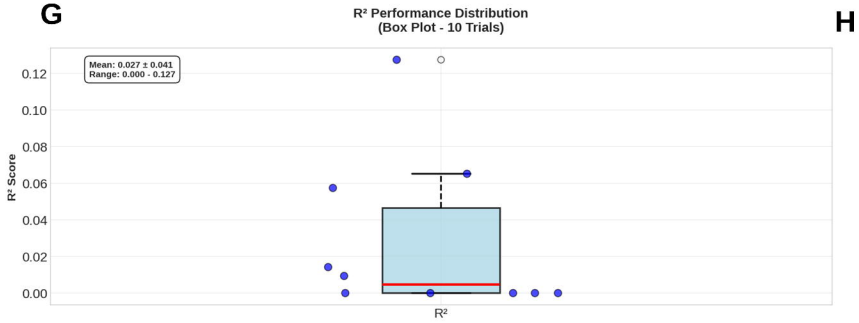
F

Statistical results summary
Performance (n=10)

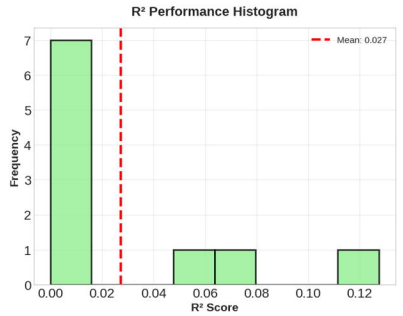
- R^2 : 0.812 ± 0.035
- 95 % CI : [0.789, 0.830]
- Spearman p : 0.911 ± 0.013
- 95 % CI : [0.902, 0.919]

BoNT/B1-SYT2 System

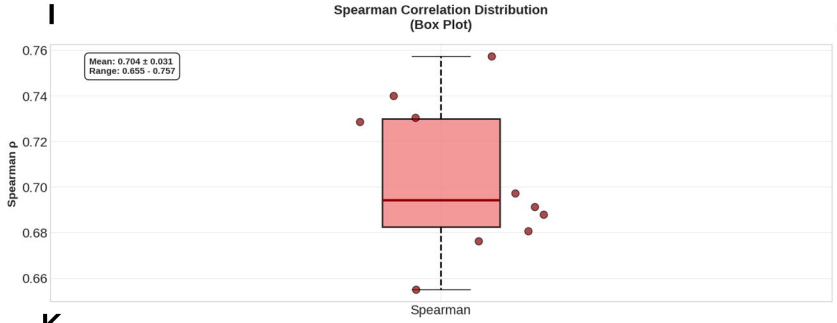
G



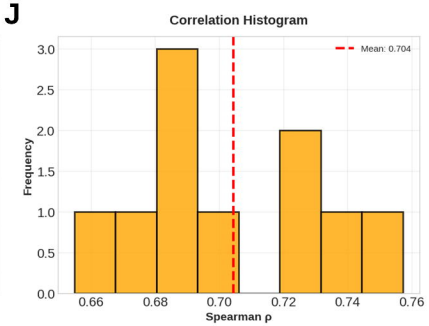
H



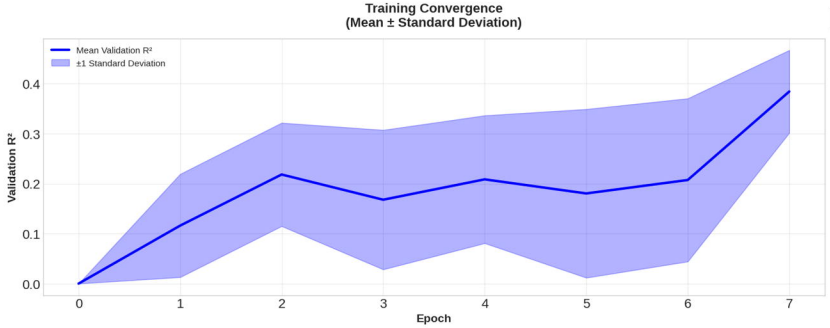
I



J



K



L

Statistical results summary
Performance (n=10)

- R^2 : 0.027 ± 0.041
- 95 % CI : [0.005, 0.053]
- Spearman p : 0.704 ± 0.031
- 95 % CI : [0.686, 0.723]