AI-Accelerated Mapping of RNA-Targeted Small-Molecule Space: Library Design, Multi-Modal Screening, and Synthesis-Aware Optimization

This review maps an AI-accelerated roadmap for discovering RNA-targeted small molecules, using **OTAVA's RNA-binding library** as the starting point. We outline design principles for building libraries enriched in RNA-privileged chemotypes; detail a multi-modal screening cascade that prioritizes solution-phase discovery with rigorous orthogonal biophysical validation; and demonstrate how synthesis-aware, AI-guided optimization compresses design–make–test–analyze cycles to speed SAR and de-risk development. Together, these practices convert focused libraries—exemplified by OTAVA's screening-ready collection—into mechanism-informed leads with realistic routes to scale-up and potential for therapeutic development.

1. Rationale for Targeting RNA with Small Molecules

RNA has emerged as a therapeutically tractable target class with roles across oncology, virology, and neurological disease. Despite challenges associated with conformational plasticity and fewer "classic" binding pockets, discrete secondary and tertiary motifs such as bulges, internal loops, junctions, riboswitch elements, and G-quadruplexes can be engaged selectively by small molecules.

2. Library Design & Preparation

A purpose-built RNA-focused library should enrich chemotypes predisposed to engage structured RNA motifs while maintaining drug-like properties and assay compatibility. The goal is to raise the prior probability of discovering selective binders for bulges, internal loops, junctions, riboswitch elements, and related architectures—then ensure those chemotypes can progress through screening, validation, and SAR without avoidable liabilities (Thomas & Hergenrother, 2008; Childs-Disney et al., 2022).

- **RNA-privileged features**
  Enrich for planarity/aromaticity to support π–π stacking with nucleobases, position H-bond donors/acceptors to address non-canonical pairs, and include cationic/polar elements to complement the polyanionic backbone (Thomas & Hergenrother, 2008; Childs-Disney et al., 2022).
  Implementation: Use substructure/descriptor filters (e.g., fused heteroaromatics, benzimidazole-like motifs) and exclude patterns linked to nonspecific or intercalative promiscuity, keeping protonation states relevant to physiological conditions (Thomas & Hergenrother, 2008; Childs-Disney et al., 2022).
- **Balanced drug-likeness**
  Keep physicochemical properties within ranges compatible with permeability and solubility,

enabling cellular follow-up and medicinal chemistry tractability (Thomas & Hergenrother, 2008).
Implementation: Apply cut-offs for MW, cLogP, pKa, HBD/HBA, and PSA with flexibility around known RNA-binder profiles; deprioritize extremes that drive colloidal aggregation or insolubility (Thomas & Hergenrother, 2008).

- **Scaffold diversity**
  Combine rigid heteroaromatics, flexible linkers, and fragment-like cores to span distinct RNA secondary/tertiary motifs and allow fragment-to-lead growth (Bernat & Disney, 2015; Wicks et al., 2023).
  Implementation: Use clustering on Bemis–Murcko frameworks and 3D shape to avoid redundancy; seed with fragment-like matter to permit efficient growth toward motif-matched vectors (Bernat & Disney, 2015; Wicks et al., 2023).

- **Benchmarking controls**
  Include positive controls (established RNA binders for the target class) and negative/orthogonal controls (inert analogs, sequence/structure variants) to calibrate assay behavior and triage artifacts (Hargrove, 2021; Chen & Miller, 2020).
  Implementation: Plate controls on every assay plate; incorporate counterscreens for dye binding/inner-filter effects and colloids; set predefined QC gates for plate acceptance (Hargrove, 2021; Chen & Miller, 2020).

- **Screening-ready formatting and QC**
  Prepare compounds as 10 mM DMSO stocks in 96- or 384-well 2D-barcoded plates with full identity/purity verification (LC–MS or ^1H NMR) to enable plug-and-play HTS/biophysics (OTAVA RNA-binding library page; Chen & Miller, 2020).
  Implementation: Provide SDF/SMILES, plate maps, and concentration certificates; minimize freeze–thaw cycles and verify buffer-compatibility/solubility for the primary assay (OTAVA RNA-binding library page; Chen & Miller, 2020).

2.1 Practical Handling and Pre-Screen Checks

Suggested good practices for handling and pre-screen evaluation

- Plate handling. DMSO stocks should be sealed and aliquoted to limit freeze–thaw cycles, moisture ingress, and evaporation.
- Assay compatibility. Solubility and short-term stability should be verified in the intended assay buffer after dilution from DMSO.
- Aggregation liability. Aggregation should be assessed (e.g., inclusion of low-level nonionic detergent, dynamic light scattering, or colloid-sensitive counterscreens).
- Optical interference (fluorescence assays). Inner-filter effects, autofluorescence, quenching, and spectral overlap should be evaluated using dye-only, ligand-only, and buffer controls.
- Plate/material effects. Nonspecific adsorption and plate-dependent biases should be checked; plate materials and surfactants should be standardized across campaigns.

Applied together, these measures reduce pan-assay artifacts and improve the interpretability and reproducibility of screening data.

3. Preparation Modalities and 2020–2025 Innovations

Over the past five years, preparation strategies for RNA-targeted libraries have moved from ad hoc curation toward systematic, modality-aware workflows that better reflect the physicochemical demands of RNA recognition and the realities of downstream screening.

- Curated physical libraries (computational enrichment)
  Curated sets are increasingly assembled by computational enrichment of vendor collections to bias toward RNA-relevant property distributions and scaffold diversity, improving *a priori* hit probability (Wicks et al., 2023; Morishita, 2022).
- Fragment-first strategies (RNA-optimized, fluorinated fragments)
  RNA-optimized, often fluorinated fragment sets enable sensitive ^{19}F NMR and complementary biophysical readouts, shortening fragment-to-lead cycles while maintaining motif selectivity (Lundquist et al., 2025; Tomemori et al., 2025).
- DNA-encoded library (DEL) adaptations for structured RNAs
  Engineering solutions that minimize nonspecific ssDNA–RNA interactions and novel architectures (e.g., "DEL Zipper") expand DEL screening to structured RNAs without sacrificing target relevance (Ma et al., 2024; Morishita, 2022).
- Assay-compatible derivatization (format-aware implementation)
  Biotin, fluorophores, or DNA barcodes are placed distal to putative RNA-binding elements, preserving native contacts while supporting AS-MS, pull-downs, and microarrays (Chen & Miller, 2020; Suresh & Disney, 2021).
- ML-guided prioritization (pre-screen triage)
  Pretrained models plus active learning now routinely rank subsets prior to wet screening, reducing experimental burden while maintaining discovery rate and rapidly incorporating early SAR feedback (Graff et al., 2020; Cao et al., 2023).

Trend summary. Collectively, these modalities illustrate a field moving from generalized small-molecule collections toward RNA-aware, screening-ready libraries that integrate computational curation, fragment and DEL innovations, assay-savvy chemistry, and data-driven triage.

4. Tiered Screening & Validation Cascade

A multi-modal cascade improves hit fidelity and accelerates progression:

Tier	Method	Advantages / Suitability	Limitations / Caveats	Best Use Case
Primary / medium throughput	Affinity Selection Mass Spectrometry/ Automated Ligand Identification System (ASMS/ALIS)	Solution-phase, label-free discovery; avoids immobilization/dye artifacts (Suresh & Disney, 2021; Momentum Bio, 2024)	Requires sensitive MS; careful buffer control and nonspecific binding suppression	General primary screen for enriched RNA-binding libraries

Tier	Method	Advantages / Suitability	Limitations / Caveats	Best Use Case
	Fluorescence assays (FP, FRET, displacement)	High throughput; adaptable triage and counter-screens (Chen & Miller, 2020; Momentum Bio, 2024)	Dye/inner-filter effects; aggregation artifacts; requires orthogonal confirmation	Rapid triage of subsets; follow-up on ASMS hits
	Small Molecule Microarrays (SMM)	Parallel profiling of chemotype–RNA interactions (Childs-Disney et al., 2022)	Surface/immobilization artifacts; optimization required	Scanning binding preferences across scaffolds
Secondary / orthogonal	SPR / BLI	Label-free kinetics (kon/koff), specificity ranking (Hargrove, 2021)	RNA immobilization chemistry can perturb structure	Orthogonal confirmation and kinetic ranking
	Isothermal Titration Calorimetry (ITC)	Gold-standard thermodynamics (Kd, ΔH, ΔS) (Hargrove, 2021)	Higher material demand; typically μM binders	Final validation of prioritized hits
	Native Mass Spectrometry (nMS)	Direct observation of noncovalent RNA–ligand complexes; stoichiometry (Fitzgerald et al., 2023)	Low-salt buffers; potential gas-phase dissociation	Validating complex formation/selectivity
Structural mapping	Chemical probing/footprinting (e.g., SHAPE, DMS)	Detects ligand-induced RNA conformational changes (Childs-Disney et al., 2022)	Signal can be subtle; workflow complexity	Binding-site localization; SAR guidance
Functional / cellular	Live-cell reporters (e.g., RiPCA, RNA sensors)	Mechanistic confirmation in cellular context (Mullangi et al., 2021; Stanford OTL, 2023)	Reporter engineering; indirect effects possible; stringent controls required	Lead confirmation in physiologically relevant systems

5. End-to-End Workflow

A coherent discovery workflow for RNA-targeted small molecules should

I. prioritize assay-ready chemotypes before wet screening,
II. emphasize solution-phase primary discovery to preserve native RNA conformations,
III. deploy orthogonal biophysical methods to confirm true binding and quantify energetics/kinetics, and

IV. integrate structural mapping and cellular readouts to establish mechanism and translate activity into relevant systems.

The sequence below reflects these principles and is grounded in methods that have proven effective across RNA-focused campaigns.

1. Virtual pre-filtering (triage before the bench)
   What & why. Motif- or structure-informed docking and ML-based prioritization are used to down-select from the library into assay-sized subsets with higher prior probability of success (Graff et al., 2020; Cao et al., 2023).
   How to execute. Define the target RNA motif(s) and negative controls (e.g., abundant off-target motifs), generate pose/score ensembles, and combine with ML ranking (active learning where available). Deliverables include a tractable “design/test pack” plus decoys for early specificity checks.

2. Primary discovery (solution-phase, label-free)
   What & why. Affinity-selection mass spectrometry (ASMS/ALIS) performed in near-physiological buffers enables label-free detection of RNA–ligand complexes and reduces immobilization/dye artifacts (Suresh & Disney, 2021; Momentum Bio, 2024).
   How to execute. Optimize ionic strength/Mg^{2+} and temperature to preserve the target fold; include sequence/structure controls; and implement competition conditions to challenge nonspecific binders. Output is a ranked list of primary binders with preliminary stoichiometry.

3. Parallel fluorescence triage (throughput with controls)
   What & why. Fluorescence polarization (FP), FRET, or displacement assays provide higher throughput confirmation and IC_{50} ranking but are susceptible to optical and colloidal artifacts; rigorous counter-screens are therefore essential (Chen & Miller, 2020; Momentum Bio, 2024).
   How to execute. Run dye-only/ligand-only controls, inner-filter/autofluorescence checks, and low-level detergent or orthogonal colloid counterscreens. Use these assays to filter ASMS actives and to rank chemotypes for biophysical follow-up.

4. Orthogonal biophysics (mechanism and quantitation)
   What & why. Surface plasmon resonance (SPR) or biolayer interferometry (BLI) provide kinetics (k_{on}, k_{off}) and help discriminate specific binding from sticky interactions; isothermal titration calorimetry (ITC) affords thermodynamics (Kd, ΔH, ΔS); native MS (nMS) confirms stoichiometry under gentle conditions (Hargrove, 2021; Fitzgerald et al., 2023).
   How to execute. Use multiple immobilization chemistries for SPR/BLI to check for surface artifacts; titrate RNA/ligand to appropriate windows for ITC; and optimize low-salt buffers for nMS. Convergent evidence across these methods increases confidence in true, direct binding.

5. Structural mapping (binding-site localization and SAR guidance)
   What & why. Chemical probing (e.g., SHAPE, DMS) or enzymatic footprinting maps ligand-induced conformational changes, localizes the binding region, and generates hypotheses that directly inform analog design (Childs-Disney et al., 2022).
   How to execute. Compare protected/perturbed nucleotides with the known secondary/tertiary model; integrate with docking/ML hypotheses to propose targeted modifications (e.g., H-bond vectors, planarity/cationicity adjustments).

6. Cellular validation (function in context)
   What & why. Engineered reporters and assays that monitor RNA–protein interactions or RNA processing provide functional confirmation in cells, ensuring that biochemical binding translates into mechanism-relevant activity (Mullangi et al., 2021; Stanford OTL, 2023).
   How to execute. Select reporters aligned with the intended mechanism (e.g., interaction disruption, splicing modulation, stability). Include orthogonal controls (inactive analogs, unrelated RNA targets) and basic developability checks (solubility, permeability proxies).

Outcome. Executed as a loop, this cascade advances hits from probabilistically enriched triage to biophysically validated binders, mapped mechanisms, and cell-active leads, while continuously informing design choices and resource allocation through data returned at each stage (Graff et al., 2020; Cao et al., 2023; Suresh & Disney, 2021; Momentum Bio, 2024; Chen & Miller, 2020; Hargrove, 2021; Fitzgerald et al., 2023; Childs-Disney et al., 2022; Mullangi et al., 2021; Stanford OTL, 2023).

6. AI-Driven Optimization & REAXENSE Integration

Access to REAXENSE platform enables an AI-accelerated DMTA loop—from enumerating design hypotheses to selecting make-ready analogs that balance potency, selectivity, and developability while remaining synthetically tractable.

Core capabilities

- Analog expansion & generative design: Proposes chemotype-focused analogs (scaffold hops, MMPs, generative ideas) while preserving key RNA-binding features (planarity, tuned cationicity, H-bond vectors).
- Multi-parameter scoring: Ranks candidates by potency/selectivity, permeability/solubility, and ADMET risks to maintain medicinal-chemistry realism.
- Synthesis-aware feasibility: Filters designs by routeability, building-block availability, step count, and throughput so proposals map cleanly to OTAVA re-supply and custom analoging.
- Active learning: Uses model uncertainty to pick high-information batches, increasing sample efficiency as SAR accumulates (Graff et al., 2020; Cao et al., 2023).
- Selectivity & assay awareness: In-silico counterscreens bias for the intended RNA motif and flag fluorescence/aggregation liabilities so sets are screen-ready (Morishita, 2022).

Workflow integration

1. Input: Seed hits + assay context and any structural/footprinting insights.
2. Design pack: Ranked 24–96 analogs with predicted profiles and make-lists.
3. Make & test: OTAVA synthesizes and plates; results feed back to the model.
4. Learn & iterate: Next batch balances exploitation (on-chemotype potency) and exploration (diversity for selectivity/developability), compressing cycles.

By coupling AI-guided design with synthesis-aware feasibility and OTAVA's execution, programs advance from hits to mechanism-informed leads with shorter cycles and fewer dead-ends (Childs-Disney et al., 2022; Graff et al., 2020; Cao et al., 2023; Morishita, 2022).

7. Formatting, QC, and Logistics

Standard delivery is provided as 10 mM DMSO stocks in 96- or 384-well 2D-barcoded plates, with custom plates or tubes, user-defined concentrations/volumes, plate maps, and controls available. Alternative formats (e.g., dry film/solids) can be supplied on request. A complete digital package (SDF/SMILES, plate maps, barcodes) is included with each shipment.

Each compound is accompanied by QC via LC–MS or ^{1}H NMR. For confirmed hits, additional QC can be performed as appropriate, and higher-purity material can be supplied. Rapid resupply (mg to multi-gram) and custom synthesis/analog expansion aligned with SAR timelines are supported.

8. Conclusion

By uniting RNA-privileged library design with solution-phase primary discovery and orthogonal biophysical validation, then layering structural mapping and cellular readouts, programs can convert tentative binders into high-confidence candidates with defined modes of action. When this workflow is augmented by AI-driven prioritization and synthesis-aware optimization, SAR cycles tighten, false positives fall, and resources concentrate on chemotypes with genuine translational promise. The result is a disciplined path from initial hits to robust, mechanism-informed leads poised for scale-up and preclinical evaluation.

References

1. Bernat, V., & Disney, M. D. (2015). RNA structures as mediators of neurological diseases and as drug targets. *Neuron*, *87*(1), 28–46. https://doi.org/10.1016/j.neuron.2015.06.012
2. Thomas, J. R., & Hergenrother, P. J. (2008). Targeting RNA with small molecules. *Chemical Reviews*, *108*(4), 1171–1224. https://doi.org/10.1021/cr0681546
3. Chen, J. L., & Miller, B. G. (2020). Tools for studying small molecule–RNA interactions in vitro and in cells. *Biopolymers*, *112*(3), e23391. https://doi.org/10.1002/bip.23391
4. Childs-Disney, J. L., Yang, X., Gibaut, Q. M. R., Tong, Y., Batey, R. T., & Disney, M. D. (2022). Targeting RNA structures with small molecules. *Nature Reviews Drug Discovery*, *21*(10), 736–762. https://doi.org/10.1038/s41573-022-00521-4
5. Fitzgerald, K., Ragunathan, K., & Znosko, B. M. (2025). Using native mass spectrometry to study RNA–ligand interactions. *bioRxiv*. https://doi.org/10.1101/2025.02.09.637332
6. Hargrove, A. E. (2021). Small molecule–RNA targeting: Starting with the fundamentals. *Trends in Pharmacological Sciences*, *42*(9), 744–757. https://doi.org/10.1016/j.tips.2021.06.006
7. Rosenblum, S. L., Allen, T. E. H., Wicks, S. L., et al. (2024). Live-cell screening to identify RNA-binding small-molecule inhibitors of the pre-let-7–LIN28 RNA–protein interaction. *ACS Chemical Biology*, *19*(3), 659–671. https://doi.org/10.1021/acschembio.3c00877
8. Stanford Office of Technology Licensing. (2023). *Sensors for high-throughput screening of RNA-modulating drugs*. *Stanford TechFinder*.

https://techfinder.stanford.edu/technology/sensors-high-throughput-screening-rna-modulating-drugs

9. Suresh, B. M., & Disney, M. D. (2021). Affinity-based methods for the discovery of RNA-binding small molecules. *ACS Chemical Biology*, *16*(2), 219–232. https://doi.org/10.1021/acschembio.0c00844
10. Wicks, S. L., Morgan, B. S., Wilson, A. W., & Hargrove, A. E. (2023). Probing bioactive chemical space to discover RNA-targeted small molecules. *bioRxiv*. https://doi.org/10.1101/2023.07.31.551350
11. Hargrove, A. E. (2020). Small molecule–RNA targeting: Starting with the fundamentals. *Chemical Communications*, *56*(93), 14788–14797. https://doi.org/10.1039/D0CC06796B
12. Ma, Z., Su, H., Yang, Y., et al. (2025). Development of a DNA-encoded library screening method “DEL Zipper” to empower the study of RNA-targeted chemical matter. *SLAS Discovery*, *31*, 100204. https://doi.org/10.1016/j.slasd.2024.100204 (Epub 2024 Dec 21)
13. Lundquist, K. P., Romeo, I., Puglielli, R. B., et al. (2025). Design, synthesis, and screening of an RNA-optimized fluorinated fragment library. *SLAS Discovery*, *31*, 100215. https://doi.org/10.1016/j.slasd.2025.100215
14. Nickbarg, E. B., Spencer, K. B., Mortison, J. D., & Lee, J. T. (2023). Targeting RNA with small molecules: Lessons learned from Xist RNA. *RNA*, *29*(4), 463–472. https://doi.org/10.1261/rna.079523.122
15. Martin, W. J., Grandi, P., & Marcia, M. (2021). Screening strategies for identifying RNA- and ribonucleoprotein-targeted compounds. *Trends in Pharmacological Sciences*, *42*(9), 758–771. https://doi.org/10.1016/j.tips.2021.06.001
16. Morishita, E. C. (2022). Discovery of RNA-targeted small molecules through merging experimental and computational technologies. *Expert Opinion on Drug Discovery*, *17*(12), 1289–1307. https://doi.org/10.1080/17460441.2022.2134852
17. Kovachka, S., Panosetti, M., Grimaldi, B., Azoulay, S., Di Giorgio, A., & Duca, M. (2024). Small molecule approaches to targeting RNA. *Nature Reviews Chemistry*, *8*(2), 120–135. https://doi.org/10.1038/s41570-023-00569-9
18. Graff, D. E., Shakhnovich, E. I., & Coley, C. W. (2020). Accelerating high-throughput virtual screening through molecular pool-based active learning. *Chemical Science*, *11*(47), 12028–12036. https://doi.org/10.1039/D0SC06805E
19. Cao, Z., Sciabola, S., & Wang, Y. (2023). *Large-scale pretraining improves sample efficiency of active-learning–based molecule virtual screening*. *arXiv preprint*. http://arxiv.org/pdf/2309.11687v1
20. Tomemori, M., Ichijo, R., Shinohara, Y., Hatta, K., Nakatani, K., & Kawai, G. (2025). NMR analysis of interactions between RNA structure elements and small molecules. *Journal of Biochemistry*, *178*(1), 1–9. https://doi.org/10.1093/jb/mvaf020

21. Momentum Bio. (2024). *Screening strategies for RNA-targeted small molecule drugs*. *Momentum Bio*. https://momentum.bio/content/screening-strategies-for-rna-targeted-small-molecule-drugs/

22. OTAVA Ltd. (n.d.). *RNA-binding library*. *OTAVA Ltd.*. https://www.otavachemicals.com/products/targeted-libraries-and-focused-libraries/rna-binding