Nucleic Acids Research, 2025, **53**, gkaf166
https://doi.org/10.1093/nar/gkaf166
Methods

In silico λ-dynamics predicts protein binding specificities to modified RNAs

Murphy Angelo[1], **Wen Zhang**[1,2,*], **Jonah Z. Vilseck**[1,3,*], **Scott T. Aoki**[1,2,*]

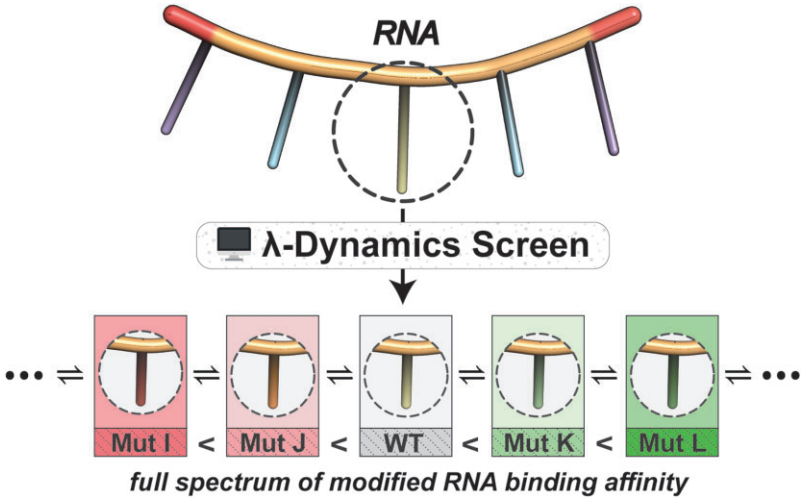
[1]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, IN 46202, United States
[2]Melvin and Bren Simon Cancer Center, 535 Barnhill Drive, Indianapolis, IN 46202, United States
[3]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN 46202, United States

[*]To whom correspondence should be addressed. Email: staoki@iu.edu
Correspondence may also be addressed to Wen Zhang. Email: wz15@iu.edu
Correspondence may also be addressed to Jonah Z. Vilseck. Email: jvilseck@iu.edu

Abstract

RNA modifications shape gene expression through a variety of chemical changes to canonical RNA bases. Although numbering in the hundreds, only a few RNA modifications are well characterized, in part due to the absence of methods to identify modification sites. Antibodies remain a common tool to identify modified RNA and infer modification sites through straightforward applications. However, specificity issues can result in off-target binding and confound conclusions. This work utilizes *in silico* λ-dynamics to efficiently estimate binding free energy differences of modification-targeting antibodies between a variety of naturally occurring RNA modifications. Crystal structures of inosine and N6-methyladenosine (m^6A) targeting antibodies bound to their modified ribonucleosides were determined and served as structural starting points. λ-Dynamics was utilized to predict RNA modifications that permit or inhibit binding to these antibodies. *In vitro* RNA-antibody binding assays supported the accuracy of these *in silico* results. High agreement between experimental and computed binding propensities demonstrated that λ-dynamics can serve as a predictive screen for antibody specificity against libraries of RNA modifications. More importantly, this strategy is an innovative way to elucidate how hundreds of known RNA modifications interact with biological molecules without the limitations imposed by *in vitro* or *in vivo* methodologies.

Graphical abstract



Introduction

Biology has an RNA complexity problem. Cells must make sense of a vast sea of RNAs that function as protein code, regulatory molecules, enzymes, scaffolds, and other biological tools. Furthermore, the four canonical RNA bases can be enzymatically modified into new chemical structures that change their ability to base pair, form secondary structure, and interact with RNA-binding proteins [1]. These chemical additions can be as small as a single methyl group or as large as a sugar moiety. Over 170 RNA modifications have been

Received: January 27, 2024. **Revised:** February 19, 2025. **Editorial Decision:** February 19, 2025. **Accepted:** February 20, 2025

identified across all three kingdoms of life [1]. Comprehensive databases are available that provide current information on RNA modification chemical structures, biosynthetic pathways, links to human disease, relevant publications, and more [2]. RNA modifications are prevalent in biology and function as an epigenetic code to regulate development [3], respond to infectious diseases [4], and are involved in cancer progression [5]. Their combinatorial complexity in nucleic acid also highlights how individual or collections of RNA modifications may alter an RNA's fate or function. A current challenge is the development of methods to identify all modification sites to decipher the roles of these RNA modifications in biology.

A variety of methods can identify a few RNA modification sites. For example, chemical treatment can identify m^6A (e.g. GLORI [6]) and pseudouridine (e.g. pseudo-seq [7]) by taking advantage of chemistries that affect a modified base differently than an unmodified base. Direct RNA nanopore sequencing can also identify specific modifications like m^6A [8–18] through differences in electrical current perturbations as the modified RNA transverses the sequencing pore. Both strategies, however, require tailor-made approaches to accommodate each RNA modification's unique biochemical characteristics. Furthermore, without employing enrichment strategies, low abundance modifications remain difficult to detect. Adaptable methods are needed to elucidate the full breadth of modified RNAs found in living organisms.

A common, versatile identification strategy uses antibodies to immunoprecipitate modified RNAs [19]. These enriched RNAs are then sequenced to identify RNA targets and infer modification sites. Immunoprecipitation and sequencing methods are well established with straightforward workflows, and enrichment permits identification of less prevalent modification sites. Indeed, much of the work determining the modification sites of N6-methyladenosine (m^6A, e.g. [20, 21]), N1-methyladenosine (m^1A, e.g. [22–25]), 5-methylcytosine (m^5C, e.g. [26, 27]), and others have relied on antibodies.

Antibodies can become *de novo* RNA-binding proteins through adaptive immunity. Immunoglobulin G (IgG) antibodies are comprised of two heavy and two light polypeptide chains that assemble a pair of six hypervariable complementary-determining region (CDR) loops at their antigen recognition interface [28–30]. Antibodies recognize a variety of antigens through CDRs that vary in amino acid length and composition. How antibodies recognize proteins is well studied [31], but how antibodies recognize modified RNAs is less clear. A polyinosine-antibody crystal structure was determined bound to various nucleotides [32]. Closer inspection of the structure reveals a large, suitably configured pocket adjacent to the bound nucleotide (Supplementary Fig. S1), suggesting that the antibody may have specificity toward nucleic acid, not single bases. Regardless, the lack of antibody structures targeting other modified bases limits insights into how antibodies recognize RNA modifications.

The success of using antibodies for RNA modification site identification depends on the quality of the antibody [33, 34]. Antibodies with low specificity have assigned erroneous biochemical functions to RNA modifications. For example, published studies reached differing conclusions regarding the mechanism of the m^1A modification. Two studies found m^1A prevalent in the 5′ ends of mRNA [24, 25], suggesting that the modification enhances translation [25], while contrasting studies reported it as rare in mRNA [22, 23]. In the former studies, it was later discovered that the antibody used for m^1A RNA enrichment also had affinity toward 7-methylguanosine (m^7G, [22]), an abundant mRNA 5′ cap modification crucial for cap-dependent translation [35]. These false positive site identifications led to incorrect conclusions regarding m^1A function. Because the identification of RNA targets and their specific modification sites gives insight into their biological and biochemical mechanisms, the development of antibodies with high affinity and high specificity is a key to successfully discovering the biological roles of the many RNA modifications. However, given the large number of RNA modifications and the subtle chemical differences between them, off-targets of RNA modification antibodies will be a continuous, inevitable problem. The current state of RNA chemistry prevents *in vitro* testing of all known RNA modifications, and thus new methods are required to predict the specificity of RNA modification-targeting antibodies.

Computational approaches have the potential to screen antibodies for their predicted ability to bind modified RNA. Physics-based, free energy calculations are an accurate, rigorous, and cost-effective means to quantify chemical probe interactions with protein structures *in silico* [36–38]. These calculations compute relative binding free energies ($\Delta\Delta G_{\text{bind}}$) between two or more molecules by transforming between alternate chemical groups. Because they also incorporate molecular dynamics (MDs) simulations, these calculations can provide dynamic structural characterization of macromolecular complexes. With these methods, changes in RNA-protein binding affinities can be monitored as a function of the chemical differences between modified or unmodified RNAs. Hence, modeling different RNA modifications can predict binding selectivity.

λ-Dynamics is an efficient, physics-based free energy method capable of accurately and rapidly comparing the relative binding affinities of hundreds of chemical modifications to a system. This method holds a key advantage over other *in silico* strategies in that it can model multiple chemical variations simultaneously within a single simulation [39, 40], making it more efficient and higher throughput. In a λ-dynamics calculation, a variable λ parameter allows chemical groups to dynamically scale between "on" and "off" states during a MD simulation. Akin to selection in an *in vitro* competitive binding assay, this dynamic behavior effectively differentiates the varying affinities of target molecules, providing insights into their binding characteristics. Thus, λ-dynamics can rapidly screen for the best binders from a library of chemical modifications [41, 42]. To date, λ-dynamics has accurately measured the relative binding free energy differences of large chemical inhibitor libraries targeting HIV reverse transcriptase [43–45] and β-secretase 1 enzymes [46, 47], of mutations at various protein-protein interfaces [48, 49], as well as of the folding free energies of a mutant T4 lysozyme enzyme [50]. Notably, chemical probe binding studies with λ-dynamics demonstrated 8- to 30-fold efficiency gains over other conventional free energy methods [43, 46]. This results in months of saved computational time, facilitating rapid binding comparisons for a large and diverse molecular library.

The following investigation tested whether λ-dynamics could accurately predict how RNA modifications affected RNA–protein interactions, a first for this methodology. This work determined the structures of two modified RNA-targeting antibodies bound to inosine and m^6A, revealing that these antibodies recognize their target ligands similar to other modified RNA binding proteins. The structural models per-

mitted the use of λ-dynamics to perform a computational screen of RNA base modifications bound to inosine and m^{6}A antibodies to predict their binding specificities. These *in silico* binding predictions were verified with *in vitro* binding assays. Collectively, the results demonstrate how structural biology can be combined with λ-dynamics to predict modified RNA–protein interactions without the limitations imposed by biochemical experiment methodologies.

Materials and methods

Recombinant antibodies

Commercial monoclonal antibodies targeting inosine (Diagenode) and m^{6}A (Abcam) were sequenced by Abterra Biosciences (San Diego, CA) [51–53]. Briefly, the antibodies were fragmented and submitted for MS/MS mass spectrometry. The data were then analyzed to predict the probable antibody sequence. Full-length monoclonal antibodies (mAb) and antibody fragments (Fab) were produced recombinantly in human cells by Sino Biological (Wayne, PA). Fabs were made from mAbs by papain protease digestion, Fc removal by protein A, and size exclusion chromatography. All mAbs and Fabs were shipped and stored in phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4).

Crystallography

Recombinant Fabs were concentrated to approximately 3–5 mg/ml and sitting drop crystal trays were set with an Oryx4 (Douglas Instruments; Hungerford, United Kingdom). The m^{6}A Fab was set up without and with 1 mM m^{6}A nucleoside (MedChemExpress). Crystals were observed by 4 weeks in the following conditions: 1) the inosine Fab in 50 mM Tris pH 8.3, 15% PEG 4000, 0.1 mM EDTA; 2) the m^{6}A Fab only in 20% (v/v) PEG 2K, 0.2 M MgCl2, 100 mM Tris pH 8.0; and 3) m^{6}A Fab with 1 mM m^{6}A nucleoside in 0.17 M ammonium sulfate, 25.5% (w/v) PEG 4000. The inosine and m^{6}A Fab only crystals were incubated in freezing conditions (inosine: 21% PEG 4K, 50 mM Tris pH 8.3, 0.1 mM EDTA, 20% glycerol, 0.2 mM inosine nucleoside (Sigma, I4125-1G); m^{6}A: 20% (v/v) PEG 2K, 0.2 M $MgCl_2$, 100 mM Tris pH 8.0, 5–15% (v/v) glycerol, 1 mM m^{6}A nucleoside) with addition of inosine and m^{6}A nucleoside freezing buffer for 30–60 min prior to freezing, respectively. X-ray diffraction data was collected at Lilly Research Laboratories Collaborative Access Team (LRL-CAT; Argonne National Laboratory; Argonne, IL) and ESRF ID30B (Life Sciences Collaborative Access Team (LS-CAT) operating at the European Synchrotron Radiation Facility (ESRF); Grenoble, France). Data were collected and processed by Lilly, UW-Madison Crystallography Core, and the authors. All data was indexed, merged, and scaled in XDS/Aimless [54]. Space groups were determined in XDS/pointless [54]. Model building and refinement were performed in Coot [55] and Phenix [56], respectively. In some of the inosine and m^{6}A Fab density maps, a large density was observed at the Fab antigen-binding site. The respective modified RNA nucleosides used in crystallization and in freezing modeled well into these densities (Fig. 1A and B). The final structures and merged reflection files are deposited at wwPDB (wwpdb.org; PDB IDs: 8SIP, 8TCA, 8VEV). Unmerged reflection data were deposited at Integrated Resource for Reproducibility in Macromolecular Crystallography (proteindiffraction.org).

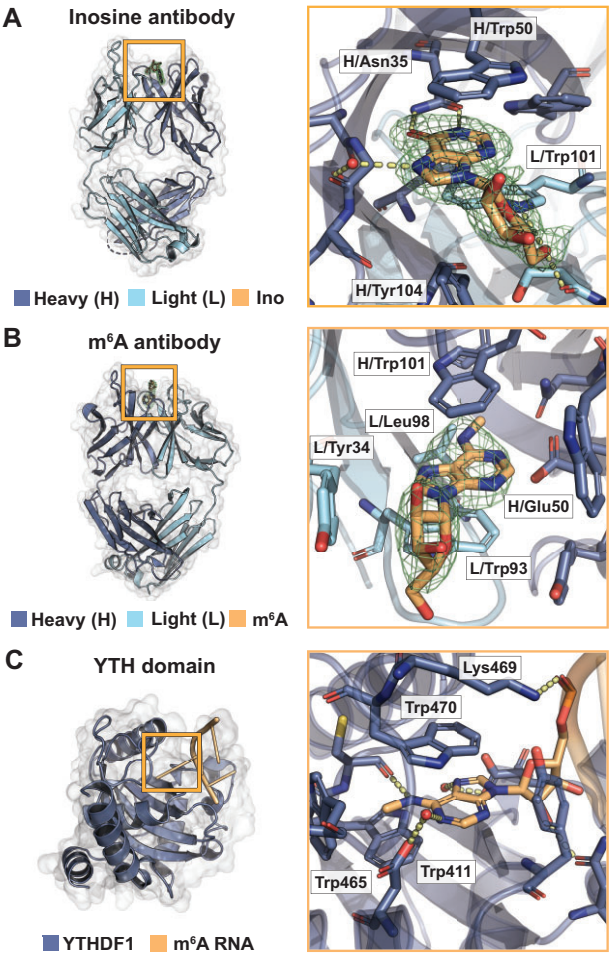


Figure 1. Binding of inosine and m^{6}A targeting antibodies mimics RNA-binding proteins. (**A**) Crystal structure of the inosine targeting antibody fragment to 1.94 Å (PDB ID: 8SIP). Overview (left) and magnified (right) rendition of the antibody bound to inosine nucleoside. $1F_oF_c$ density without ligand in mesh. Heavy chain (H), light chain (L), and inosine as labeled. Interacting amino acids include heavy chain residues Asn35, Trp40, Trp50, Gly99, Tyr104, and Leu106 and light chain residues Ser97 and Trp101. Those discussed in the main text are labeled. (**B**) Crystal structure of the m^{6}A-targeting antibody fragment to 3.06 Å (PDB ID: 8VEV). Labeling same as in (A), except m^{6}A nucleoside in orange. Interacting amino acids include heavy chain residues Trp33, Asn35, Glu50, Tyr61, Trp101, and Phe105 and light chain residues Tyr34, Trp93, and Leu98. Those discussed in the main text are labeled. (**C**) Structure of a YTH bound to m^{6}A (YTHF1, PDB ID: 4RCJ). Interacting amino acids include Tyr397, Asp401, Trp411, Cys412, Asn441, Trp465, Lys469, Trp470, and Asp507. Those discussed in the main text are labeled.

System setup for molecular modeling

Coordinates for the inosine and m^{6}A Fabs were obtained from our Protein Data Bank (PDB) entries 8SIP and 8VEV. Residue flips for His, Gln, and Asn were assessed using the MolProbity webserver [57]. Protonation states of titratable residues were assigned based on their predicted pKa values at pH 7.0 using PROPKA [58,59]. The protein–nucleoside complexes were then solvated using the CHARMM-GUI Solution Builder [60], requiring a minimum of 10 Å of solvent padding from each box face. The resulting cubic water box dimensions

were 101 Å per edge for the inosine system and 98 Å per edge for the m^6A system. Sufficient K^+ or Cl^- ions were added to neutralize the net charge of each system. Additional K^+, Mg^{2+}, and Cl^- ions were then added to achieve a final ionic strength of 150 mM KCl and 0–5 mM $MgCl_2$. This process was repeated to solvate the individual nucleosides without their respective Fabs, yielding unbound model systems with cubic box dimensions of 30 Å per edge for inosine and 32 Å per edge for m^6A.

All simulations were performed using the CHARMM molecular simulation package ([61, 62], developmental version c47a2) with the Basic λ-Dynamics Engine on graphics processing units (GPUs) [63]. Prior to running MD, each system was subjected to 250 steps of steepest descent minimization. MD simulations were then run in the isothermal-isobaric (NPT) ensemble at 25°C and 1 atm using a Langevin thermostat and Monte Carlo barostat [63–65]. The g-BAOAB integrator was used with an integration timestep of 2 fs and trajectory frames were saved every 1000 steps [63, 66]. Bond lengths between hydrogens and heavy atoms were constrained using the SHAKE algorithm [67–70]. Periodic boundary conditions were employed in conjunction with Particle Mesh Ewald (PME) electrostatics [71–73], to compute long-range electrostatic forces, and force-switched van der Waals (vdW) interactions [74]. Non-bonded cutoffs were set to 10 Å, with force switching taking effect starting at 9 Å.

All explicit solvent calculations were conducted with the TIP3P water model [75]. The CHARMM36 protein force field was used to represent the inosine and m^6A Fabs, and the CHARMM36 nucleic acid force field was used to represent the RNA oligos [76–80]. The DISU patch in CHARMM was used to correctly model all disulfide bonds in the Fabs. Modified ribonucleobase parameters were used to model non-canonical bases in the ribonucleoside [81]. For the alchemical perturbations performed with λ-dynamics, ribonucleoside base mutations were represented using a hybrid multiple-topology approach [82]. In the case of purine-to-purine mutations, analogous atoms in the shared core were harmonically restrained to one another using the Scaling of Constrained Atoms interface described previously [83].

λ-dynamics calculations

From 112 parameterized modified ribonucleobases available [81], a library of 48 bases, comprising 44 modified and 4 unmodified base candidates, were selected for *in silico* screening with λ-dynamics. Those with charged functional groups were excluded to eliminate the need to include additional correction schemes accompanying charge change perturbations [73, 84–86]. Bulky base side chains or modifications to the ribose sugar were also excluded to focus this study on modified bases that were synthetically tractable or commercially accessible to enable subsequent experimental validation. Simulations were conducted for each of the 48 ribonucleosides with λ-dynamics to alchemically transform wild-type nucleoside bases (inosine or m^6A) into a corresponding mutant base and compute relative differences in binding affinities. Prior to initiating λ-dynamics production sampling, appropriate biasing potentials must first be identified. The Adaptive Landscape Flattening (ALF) [50,87] algorithm was used to identify optimal biasing potentials to facilitate dynamic and frequent alchemical transitions between the perturbed bases. For each perturbation, ALF identified initial biases by first conducting one hundred simulations of 100 ps sampling, followed by 13 simulations of 1 ns each. These biases were then further refined via five replicate simulations of 5 ns each. With optimal biases identified, five independent production simulations of 25 ns were conducted, with an initial 5 ns of sampling removed from free energy determinations for equilibration. Run on a NVIDIA Tesla V100, these calculations took approximately 20 min/ns, enabling a single 25 ns production job to be completed in ca. 8–9 h. Ribonucleosides that unbound from the Fab binding site during λ-dynamics production sampling were labeled as unfavorable and were not pursued further. In all other cases, final $\Delta\Delta G_{bind}$ values were calculated by Boltzmann reweighting the end-state populations to the original biases with WHAM [50, 88]. Uncertainties (σ) were calculated by bootstrapping the standard deviation of the mean across each of the five independent trials. From these results, modified oligonucleotides were selected for synthesis based on commercial availability. To assess simulation convergence, five bases with a range of predicted binding affinities were selected from each Fab system for an extended *post hoc* screening. For each of these 10 bases, the optimized biases identified from the respective 25 ns simulations were utilized to conduct five additional independent production simulations of 100 ns, discarding the initial 25 ns from free energy determinations for equilibration. Convergence in the original calculations was confirmed by establishing that the root mean square error and mean unsigned error between the 25 and 100 ns predicted $\Delta\Delta G_{bind}$ values were within the statistical noise expected from the 25 ns simulation $\Delta\Delta G_{bind}$ uncertainties. λ-Dynamics input files, topology and parameter files, and initial starting structures required to replicate the calculations from this manuscript are accessible via a public repository (DOI: zenodo.org/records/13 921 976).

RNA oligonucleotide preparation

RNA oligonucleotides used for binding affinity measurements and crystallographic studies were synthesized on an ABI 394 DNA/RNA synthesizer (Applied Biosystems (ABI); Waltham, MA). m^6A (10–3005-90; Glen Research; Sterline, VA), $m^6{}_2A$ (ANP-8626; Chemgene; Wilmington, MA), and inosine (ANP-5680; Chemgene) modified RNA phosphoramidites; Biotin phosphoramidite (CLP-1517; Chemgene); and canonical RNA (A, ANP-5671; U, ANP-5674; C, ANP-6676; Chemgene) phosphoramidites were purchased from commercial sources. The canonical and modified phosphoramidites were concentrated to 0.1 M in acetonitrile. Coupling was carried out using a 5-benzylthio-1H-tetrazole (5-BTT) solution (0.25 M) as the catalyst. The coupling time was 650 s. 3% trichloroacetic acid in methylene chloride was used for the detritylation. Syntheses were performed on control pore glass (CPG-1000) immobilized with the appropriate nucleosides. All L-oligonucleotides were prepared with DMTr-on and in-house deprotected using AMA (1:1 v/v aqueous mixture of 30% w/v ammonium hydroxide and 40% w/v methylamine) for 15 min at 65°C. The RNA strands were additionally desilylated with $Et_3N{\bullet}3HF$ solution to remove TBDMS groups. The 5′-DMTr deprotection was carried out using the commercial Glen-Pak purification cartridge (Glen Research). Purification was initially performed by the commercial Glen-Pak purification cartridge, followed by further purification with a 15%

denaturing PAGE gel. The oligonucleotides were collected, lyophilized, desalted, re-dissolved in water, and then concentrated as appropriate for downstream experiments. Concentrations of the aqueous RNA samples were determined by their UV absorption at 260 nm, using the Thermo Scientific Nanodrop One Spectrophotometer. The theoretical molar extinction coefficients of these samples at 260 nm were provided by Integrated DNA Technologies. RNA oligo sequences used are reported in Supplementary Fig. S8D.

ELISA

Biotin-labeled, RNA oligos of inosine, adenosine, uridine, and cytidine to test the inosine antibody binding. Cytidine oligos with single base changes of adenosine, m^6A, and $m^6{}_2A$ were synthesized to test the m^6A antibody binding. The biotin-labeled oligos were diluted to 100 nM in ELISA blocking buffer (PBS, 0.05% Tween-20, 0.2 mg/ml bovine serum albumin (BSA, BP9706100; Fisher Scientific; Hampton, NH)), and 100 ul were incubated in clear, 96-well NeutrAvidin™ Coated Plates (PI15217; Pierce; Waltham, MA) overnight at 4°C. Three technical replicates for commercial antibodies and two technical replicates for recombinant antibodies were set for each RNA oligo. No inosine or m^6A antibody wells were used to control for secondary antibody background. ELISA blocking buffer without oligos were used as a background negative control. The plates were washed with PBS-T (PBS with 0.05% Tween-20) three times, and varying concentrations of recombinant mAb incubated in each well for 1 h at room temperature (approximately 20°C). A no-mAb condition was used as a no primary antibody control. Plates were washed three times again with PBS-T and incubated with goat anti-mouse IgG conjugated to horseradish peroxidase (HRP, NBP2-30347H; Novus Biologicals; Centennial, CO) at 0.05 μg/ml in ELISA blocking buffer for 1 h at room temperature (approximately 20°C). The plates were washed again with PBS-T and incubated with 50 ul of room temperature 1-Step Ultra TMB-ELISA Substrate Solution (PI34028; Pierce). After 15 min, the reaction was stopped with 50 ul of 2M Sulfuric Acid (A300S-500, Fisher Scientific). The plates were analyzed by 450 nm absorbance with a Synergy H1 microplate reader (BioTek Instruments; Winooski, VT). All ELISA experiments were replicated at least three times. The three cleanest runs were reported. Averages, standard deviations, and graphs were performed and made in GraphPad Prism version 10.1.1 for MacOS (GraphPad Software, Boston, MA).

MST

Antibodies were dialyzed in MST binding buffer (20 mM HEPES pH 7.5, 100 mM NaCl), and concentrated protein subjected to 13 two-fold serial dilutions for 14 samples. FAM fluorophore labeled RNA was added to the dilutions for a final concentration of 100 nM RNA and loaded into Monolith NT.115 Premium Capillaries (MO-K025, Nanotemper). The assay was performed on Monolith NT.115 (NanoTemper) with its BLUE excitation laser at a power of 40%; MST laser power of 20%, 40%, and 80%; equilibrium at 1 s; data collection at 5 s; and recovery thereafter. The data were analyzed on NTAffinity software (NanoTemper) and graphs made in Prism.

Results

Crystal structures of antibody fragments bound to their modified RNA base targets

The goal was to test whether λ-dynamics could be used as an *in silico* strategy to accurately probe modified RNA–protein interactions. Antibodies can serve as modified RNA-binding proteins. They are commonly used as reagents to enrich for modified RNAs and determine modification sites in biology [19]. Currently, RNA modification targeting antibodies are relatively few in number, have modest affinity toward their targets [33, 34], and can have specificity issues that confound biological conclusions [22]. An antibody specificity screening method for known RNA modifications will enable a comprehensive view of the RNAs enriched and provide insight into how to improve antibody design.

High-resolution structures of antibodies targeting single modified RNA bases have not been published. An inosine-targeting antibody structure is available [32], but an open pocket adjacent to the nucleoside binding site potentiates the chance that the antibody binds to a dinucleotide substrate (Supplementary Fig. S1). To avoid this confounder, additional antibody structures bound to modified ribonucleosides were pursued. *In vitro* binding assays confirmed commercial antibodies had specific binding affinity toward their modified base target (Supplementary Fig. S2). The protein sequences of these antibodies were predicted by mass spectrometry and sequencing (see Methods). Recombinant antibodies were produced in cell culture and used to generate antibody fragments (Fabs). Fabs were screened in crystallizing conditions, and crystals were soaked or grown with target nucleoside ligands (see Methods). These efforts led to the determination of three modified RNA-targeting antibody crystal structures (Supplementary Table S1): one targeting inosine at 1.94 Å and two targeting m^6A at 2.02 and 3.06 Å.

IgG antibodies are composed of heavy and light protein chains, forming 6 variable loops on each arm, or antibody binding fragment (Fab), that typically dictate binding affinity to its target substrate [28–30]. In the 1.94 Å inosine and 3.06 Å m^6A antibody structures, a large, discontinuous density was observed at these variable loop regions where a modified purine target nucleoside could be adequately modeled (Fig. 1A,B). Rather than binding to loops on the periphery, the modified nucleosides bound to a central cavity created by the 6 variable loops between the heavy and light chains (Fig. 1A,B). Binding of small molecules at this location has been observed in other antibody structures [89]. In the third 2.02 Å m^6A targeting antibody structure, density in this binding pocket was not observed (Supplementary Fig. S3A). Thus, two structures yielded high-resolution models of how purine modified bases bind to antibodies.

Small molecule antibodies are selected through adaptive immunity to target a particular hapten [90]. Thus, antibodies become RNA-binding proteins through adaptation and can inform on how biology designs a protein to bind an RNA modification de novo. Modified RNA-binding proteins provide exemplary examples of potential binding architecture. For example, the YTH domains bind to m^6A with high specificity [91]. This domain arranges its side chains to (1) create a specificity pocket for the parent base and modification, (2) bind the nucleobase through stacking, and (3) line the pocket periphery with positively charged side chains to accommodate the negatively charged RNA phosphate backbone (Fig. 1C).

This "aromatic cage" formed by aromatic and hydrophobic side chains is a hallmark of the YTH family of m6A binding proteins [92–94]. Antibodies targeting modified RNAs might also mimic this strategy. Alternatively, they might use a collection of novel binding strategies, like the hydrophobic pocket formed by IGF2BP1 [95], each selected randomly through adaptive immune selection.

The inosine and m^{6}A antibody structures both bound to their modified ribonucleoside ligands in a manner similar, but not identical, to other modified RNA-binding proteins. For example, YTH domains interact with m^{6}A by using aromatic residues to stack with the nucleobase, while an additional aromatic or leucine residue creates a hydrophobic "cage" (Supplementary Fig. S3B) [92–94]. Both antibodies used paired tryptophans to create a slot for favorable stacking and a tyrosine for ribose ring interactions (Fig. 1A and B). Of note, these tryptophans and tyrosine came from differing variable loops in each antibody and are organized differently in their central antibody binding pocket (Fig. 1A and B). The difference in binding pocket organization potentially reflects how these two antibodies were isolated from different animals with separate adaptive immune responses. To specify the modified base, the inosine targeting antibody used an asparagine to select for the O6 oxygen and N1 nitrogen of the inosine nucleobase (Fig. 1A, Supplementary Fig. S3B). The added size and charge of an adenosine N6 amine versus O6 oxygen may prevent the nucleobase from entering the antibody binding site. A tryptophan (heavy chain, W47) would sterically clash with the N2 amine of guanosine. Inosine does not have this amine and thus fits properly. The m^{6}A-targeting antibody created a hydrophobic pocket to accommodate the m^{6}A methyl group (Fig. 1B, Supplementary Fig. S3B) and a glutamate side chain to hydrogen bond with the adenosine N1 nitrogen (Fig. 1B). The hydrophobic binding pocket was formed with leucine (L98, light chain) and phenylalanine (F105, heavy chain) side chains (Supplementary Fig. S3B), deviating from the single residue assembling the hydrophobic cage in YTH proteins (Supplementary Fig. S3B) [92–94]. Interestingly, glutamate side chain coordination is also observed in some YTH domains that bind m^{6}A (Fig. 1C, [96]). In sum, the antibody-ligand structures revealed that these two antibodies use similar strategies to bind their modified base targets that may permit differentiation between unmodified base counterparts.

λ-dynamics can probe how RNA modifications affect protein binding

The quality of the structures enabled predicting *in silico* how these antibodies may interact with other RNA nucleobases. There are over 140 different RNA modifications identified in biology, many of which are not available as commercial reagents or lack protocols to synthesize *in vitro*. A library of 44 modified and 4 unmodified nucleobases was selected based on published parameters for RNA modifications in the CHARMM force field [81] and their commercial availability for experimental testing *in vitro* (Supplementary Fig. S4). λ-Dynamics was used to assess differences in relative binding free energies between inosine or m^{6}A versus each library nucleobase when bound to their respective antibodies (see "Materials and methods" and Fig. 2). During the simulations, some of the modified nucleosides unbound from the antibody (Supplementary Figs S5 and S6), presumably due to having poor binding affinity or steric clashes, and were removed from further study (Supplementary Tables S2 and S3). Similar to previously performed studies [43–45, 48–50], relative binding free energies ($\Delta\Delta G_{bind}$) were calculated for the nucleosides that remained antibody-bound. Examples of the results obtained are shown (Figs 3 and 4) with full results reported in the Supplement (Supplementary Tables S2 and S3). Simulation convergence was confirmed *post hoc* by conducting extended simulations for a representative subset of the bases with predicted $\Delta\Delta G_{bind}$ values (Supplementary Fig. S7). A positive $\Delta\Delta G_{bind}$ value indicates poorer binding and a negative value suggests enhanced binding when compared to the native inosine or m^{6}A base. As a control, inosine and m^{6}A modified bases were perturbed into an identical but distinct copy of themselves within their respective antibody complexes. These free energy differences were near zero (Figs. 3A and 4), as expected of a base replacing itself, also indicating that the λ-dynamics calculations were working correctly.

λ-Dynamics predicted differing specificities and off-targets for these two antibodies. The inosine antibody had many predicted off-targets that included uridine (Fig. 3A) and uridine modifications (Fig. 3B). Inspection of the models revealed that hydrogen bonding of the asparagine side chain to the O6 oxygen in inosine could be satisfied by the O4 oxygen in uridine (Supplementary Fig. S8A). Many uridine modifications had

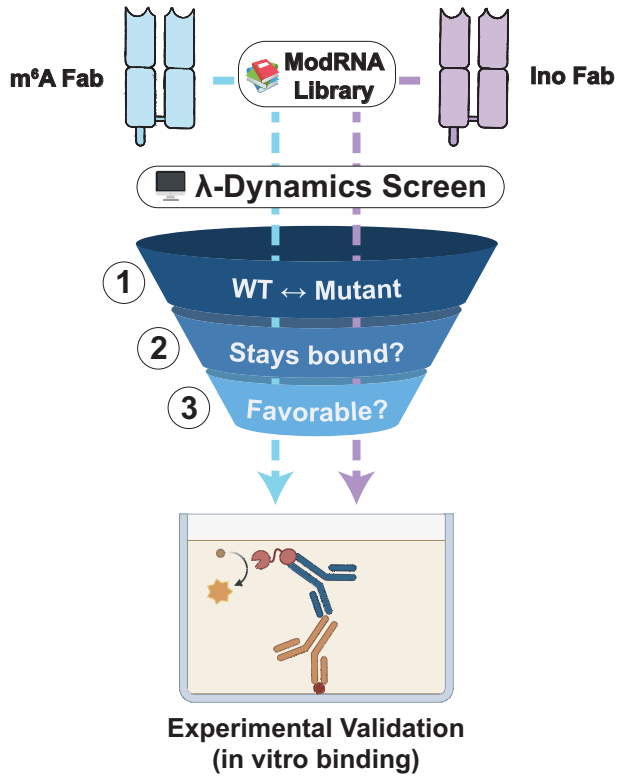


Figure 2. *In silico* λ-dynamics workflow for screening potential binders to the inosine and m^{6}A antibodies. A three-step process was used to filter candidates from a library of 48 ribonucleosides for *in vitro* antibody binding validation [1]. For each mutant library candidate, a λ-dynamics simulation was conducted to calculate a relative binding free energy between the mutant and its respective native ribonucleoside ligand (inosine or m^{6}A) [2]. All ribonucleosides that unbound during these simulations were deemed unfavorable and excluded from further processing [3]. Mutant bases with relative binding free energies deemed favorable ($\Delta\Delta G_{bind} \leq$-0.7 kcal mol^{-1}) were selected for *in vitro* validation with binding assays (ELISA) based on commercial availability.

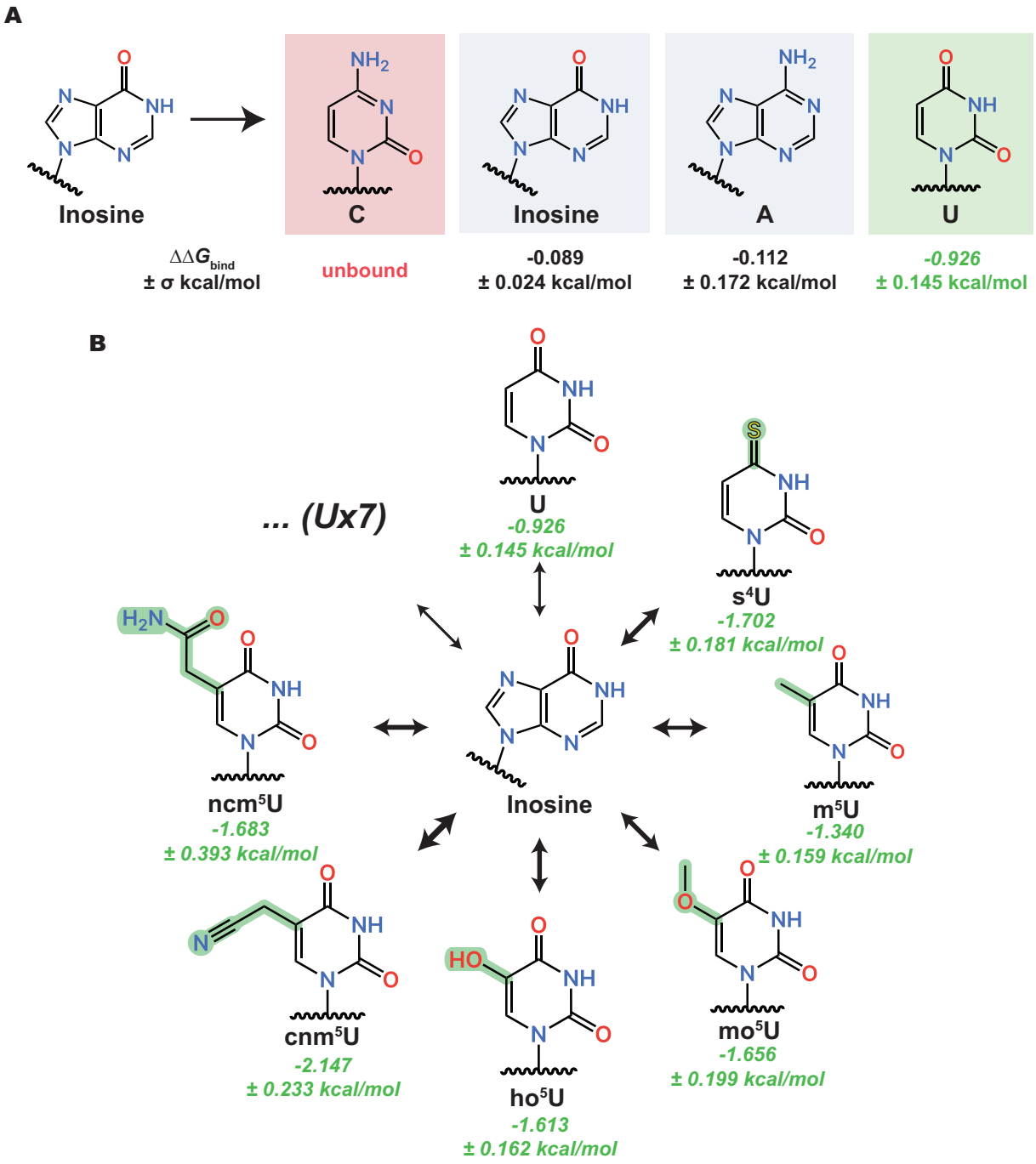


Figure 3. Highlighted binding trends from the inosine antibody λ-dynamics screen. (**A**) λ-Dynamics predicts loss of binding for cytidine (C), no change in binding for inosine and adenosine (A), and enhancement of binding for uridine (U). Estimated relative binding free energies ($\Delta\Delta G_{bind}$) and uncertainties ($\pm\sigma$) are listed. (**B**) The predicted inosine antibody promiscuity for U generalizes to many of its derivatives. Estimated relative binding free energies and uncertainties are listed. The thickness of each equilibrium arrow is proportional to the favorability of the corresponding transition. Seven other uridine derivatives (Ux7) showed enhanced binding but are not depicted. See Supplementary Table S2 for a complete list.

an O4 oxygen available for hydrogen bonding, potentially explaining why related molecules all had higher predicted binding affinities in the λ-dynamics calculations. In contrast, cytidine and adenosine were not predicted to enhance binding (Fig. 3A and Supplementary Table S2). Both nucleosides have nitrogens at similar positions, potentially making the pocket less favorable for these bases to interact by removing hydrogen bonding. Finally, a further inspection of the structures revealed a larger binding pocket in the inosine versus the m^6A antibody binding pocket (Fig. 1A,B). This larger pocket may accommodate a greater variety of shapes and sizes, increasing the propensity for off-targets. Thus, λ-dynamics predicted the inosine antibody to have many off-targets in this modestly sized ribonucleoside library.

In contrast to the inosine antibody, λ-dynamics predicted that the m^6A antibody had relatively few off-targets (Supplementary Table S3). As discussed previously, the binding pocket was small and required a N1 nitrogen on the nu-

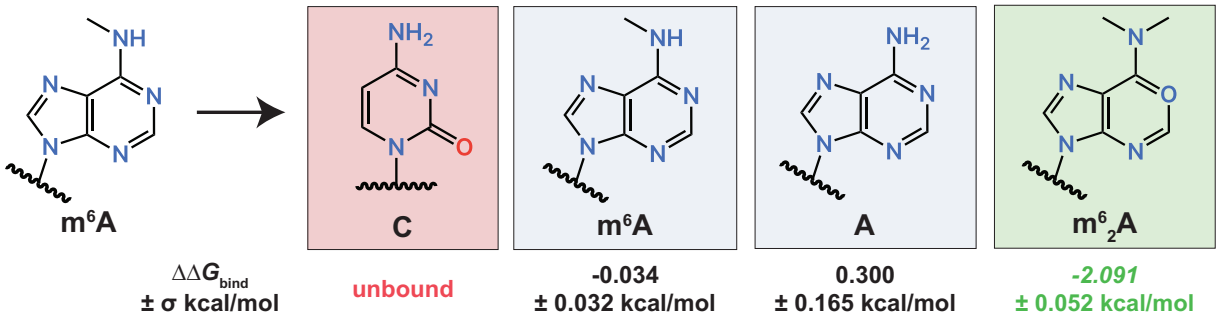


Figure 4. Highlighted binding trends from the m^{6}A antibody λ-dynamics screen. λ-Dynamics predicts loss of binding for cytidine (C), no change in binding for m^{6}A and adenosine (A), and enhancement of binding for m$^6{}_2$A. Estimated relative binding free energies ($\Delta\Delta G_{bind}$) and uncertainties ($\pm\sigma$) are listed. See Supplementary Table S3 for a complete list.

cleobase for hydrogen bonding (Fig. 1B). Along with m^{6}A, a few adenosine bases were predicted to bind (Fig. 4 and Supplementary Table S3), including adenosine and N6,N6-dimethyladenosine (m$^6{}_2$A), a dimethyl modification at the N6 nitrogen position (Supplementary Fig. S8B and C). Closer inspection of the structure revealed that the hydrophobic pocket had enough space to accommodate a second methyl group (Supplementary Fig. S8C). Similar to the inosine antibody, cytidine was predicted to be a poor binder with a high, positive free energy difference (Fig. 4). In summary, the m^{6}A antibody had fewer off-targets compared to the inosine antibody but still was predicted to bind to nucleosides other than m^{6}A.

In vitro assays support λ-dynamics' predicted binding trends

While λ-dynamics has demonstrated accuracy with modeling protein–protein and protein–small molecule binding interactions [43–49], it has so far been untested with respect to reproducing protein–RNA interactions. To evaluate our *in silico* predictions *in vitro*, Enzyme-Linked Immunosorbent Assays (ELISAs, Fig. 5A and B) and MicroScale Thermophoresis (MST, Fig. 5C and D) were used to probe the binding of inosine and m^{6}A antibodies to on- and off-target RNA bases. RNAs were synthesized using solid-state chemistry (see "Materials and methods," Supplementary Fig. S8D) to produce RNA oligomers containing inosine, adenosine, uridine, and cytidine to test inosine antibody binding, as well as m^{6}A, m$^6{}_2$A, adenosine, and cytidine to test m^{6}A antibody binding. In the ELISAs, biotin-labeled oligos were immobilized in wells, and binding of inosine and m^{6}A antibodies was assessed indirectly by detecting a secondary antibody using a chromogenic substrate. MST directly measures antibody binding affinity by using fluorophore-labeled oligonucleotides and infrared-driven diffusion [97]. Thus, two binding assays were used to assess antibody binding to a variety of RNA substrates.

The inosine and m^{6}A antibody *in vitro* binding results agreed with the λ-dynamics predictions (Fig. 5). The inosine antibody bound to inosine and uridine oligos (Fig. 5A,C), although uridine binding was only observed at much higher antibody concentrations and had lower measured affinity by ELISA and MST, respectively. This discrepancy may be due to using RNA oligomers to test *in vitro* binding, in contrast with the ribonucleoside monomers simulated with λ-dynamics, or other limitations of *in silico* modeling (see "Discussion"). In contrast, the inosine antibody did not bind to adenosine or cytosine oligos by ELISA (Fig. 5A) or have measurable affinity to cytosine by MST (Fig. 5C). The m^{6}A antibody bound to on-target m^{6}A RNA, as well as off-target m$^6{}_2$A- and adenosine-RNA in both assays (Fig. 5B and D). This differed to the results observed with the commercial m^{6}A antibody (Supplementary Fig. S2B) and may reflect antibody sequence differences between the commercial and recombinant antibodies (see "Materials and methods"). The assays also measured poor binding to cytidine only RNAs (Fig. 5B and D), as expected for a negative control. Thus, the *in vitro* binding results matched the predictions of λ-dynamics, supporting the accuracy of this *in silico* method to identify modified RNA–protein interactions.

Discussion

With hundreds of RNA modifications identified in biology, new methods are required to determine the RNA sites of each of these chemical changes to determine their functions. Antibodies targeting RNA modifications are a versatile tool to enrich and determine modification sites (reviewed in [19] and e.g. [20–27]), but their reliability hinges upon their accuracy. To this end, inosine and m^{6}A antibody structures bound to their modified ribonucleoside targets were determined to high resolution. These structures then facilitated the use of λ-dynamics, an *in silico* free energy calculation, to estimate how the antibodies may bind other unmodified and modified RNA bases, with worsened, neutral, or enhanced binding affinities. λ-Dynamics predictions matched well with *in vitro* binding assay results, supporting the accuracy of using this computational approach to measure untested RNA–protein interactions. In its simplest application and as performed in this work, the method can be used to determine off-target RNA base interactions with antibodies used for modified RNA enrichment and site identification. But the strategy holds greater promise to inject insight into the biochemical mechanisms of RNA modifications by determining how any modified RNA, commercially available for biochemical investigation or not, may interact with proteins and other molecules (Fig. 6).

The determined antibody structures targeting modified purines revealed identical binding strategies toward their respective modified RNA bases, reminiscent of modified RNA-binding proteins. Each antibody formed a specificity pocket and used tryptophans to create a slot for stacking with the nucleobase. Only one of these tryptophans had a similar sequence position between the two antibodies. The other came from a separate loop, leading to RNA binding in completely different orientations. These antibodies were created through

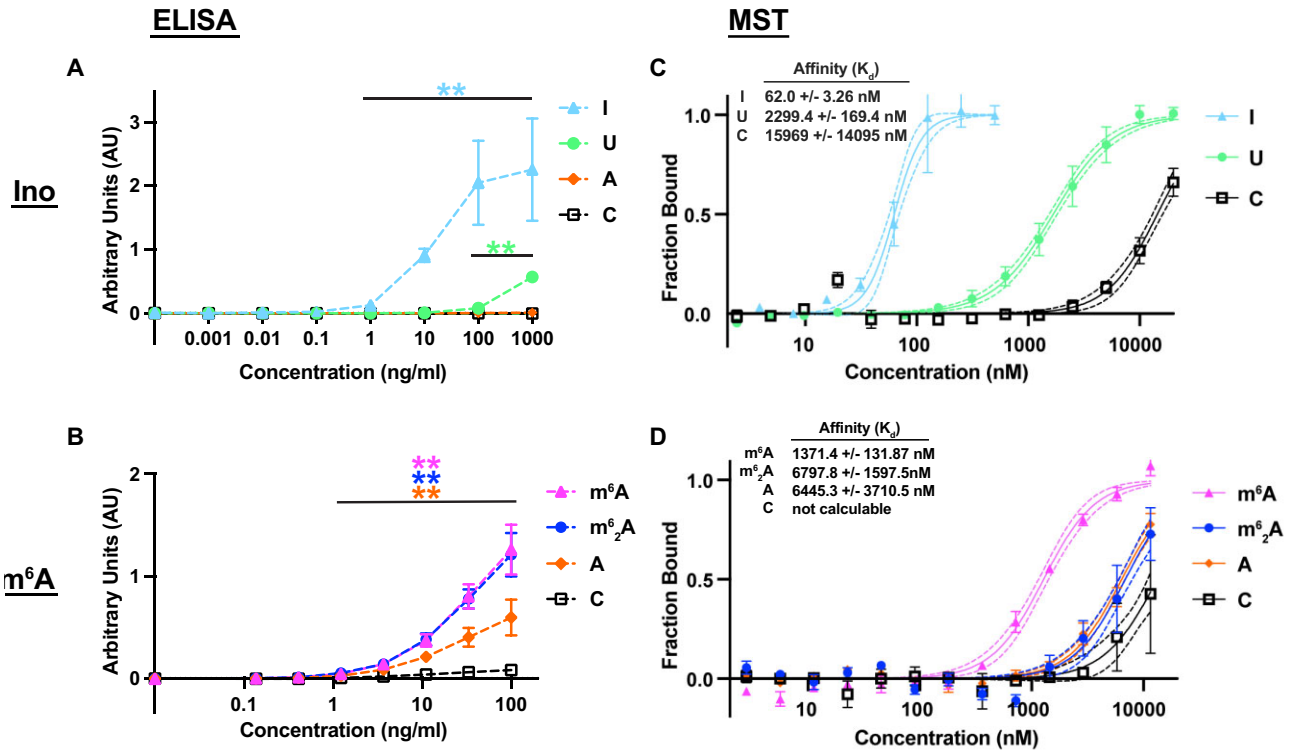


Figure 5. *In vitro* binding assay results support λ-dynamics predictions of antibody off-targets. (**A**,**B**) ELISA binding assays. (**A**) Absorbance units reported by ELISA indicating the binding affinity of inosine antibody to inosine (I), uridine (U), adenosine (A), and cytidine (C) over varying protein concentrations. Double asterisks (**) denote a *P*-value < 0.01. Inosine serves as a positive control. In line with λ-dynamics predictions, U identified as an off-target while A and C demonstrated negligible binding. (**B**) Absorbance units reported by ELISA indicating the binding affinity of m^6A antibody to m^6A, $m^6{}_2A$, adenosine (A), and cytidine (C) at varying protein concentrations. Double asterisks (**) denote a $P < 0.01$. m^6A serves as a positive control. Again, matching λ-dynamics predictions, $m^6{}_2A$ and A are identified as off-targets while C demonstrated negligible binding. All *P*-values calculated are available in Supplementary Fig. S8E and F. (**C** and **D**) MST binding assays. Dotted lines delineate 95% confidence interval (CI). (**C**) Fraction RNA bound to the inosine antibody reported over increasing protein concentration. Calculated affinities are as shown. Sum of squares f test between I versus U and C versus U curves is $P < 0.0001$. (**D**) Fraction RNA bound to the m^6A antibody reported over increasing protein concentration. Calculated affinities are as shown. Sum of squares f test between m^6A versus $m^6{}_2A$, m^6A versus A, $m^6{}_2A$ versus C, and A versus C is $P < 0.0001$; $m^6{}_2A$ versus A is $P = 0.3241$.

Figure 6. Proposed strategy to predict how proteins bind canonical and modified RNA. (1) Starting with an RNA-protein structural model, (2) an *in silico* λ-dynamics screen can be conducted to assess the favorability of the protein's interactions with a complete range of RNA bases. (3) This approach provides an economical and effective means to explore the full extent of a protein's RNA-binding capabilities that can be tested further *in vitro*.

adaptive immunity, supporting the notion that mimicking modified base RNA-binding proteins by creating a specificity pocket and using stacking for nucleobase interactions is a competent way to bind a modified nucleobase. Thus, convergent adaption may have led both purine-targeting antibodies to follow a similar binding strategy as modified RNA-binding proteins. Examples of pyrimidine-targeting antibody structures will be necessary to probe whether other modified RNA-targeting antibodies bind to their targets similarly.

Antibodies are heavily used reagents to enrich modified RNA for sequencing and site identification. This strategy has been used to identify sites of many different RNA modifications to deduce their biological and biochemical mechanisms. Regardless of new methodologies to determine RNA modification sites, antibodies will continue to be used to enrich for less abundant modifications. Thus, antibody binding to off-target RNA modifications will continue to be a problem in research. The chemical similarities between many RNA modifications make antibody specificity an expected complication. This work demonstrates how λ-dynamics is a viable *in silico* tool to determine potential RNA off-targets of antibodies. The method does not require the availability of modified nucleosides, RNA oligomers, or other *in vitro* reagents that are currently unavailable. With an accurate, high-resolution structural model, λ-dynamics can test the full breadth of RNA modifications in biology. Additionally, λ-dynamics has previously investigated the effects of protein mutations on binding [48, 49]. The method can thus be used to rationally design antibodies for improved binding specificity and affinity.

This is the first study to use λ-dynamics to probe nucleic acid–protein interactions via nucleic base perturbations. Other *in silico* molecular modeling and free energy methods have been employed to study nucleic acid–protein interactions, including predictions of DNA/RNA binding to proteins [98–100] and probing mutations in nucleic acid–protein complexes [101–103]. λ-Dynamics has several key attributes that make it advantageous over other *in silico* calculations. First, λ-dynamics enables multiple modified bases to be calculated within a single simulation. This can drastically improve efficiency over other free energy methods that can only investigate a single perturbation at a time, therefore requiring

many simulations to study multiple perturbations. Second, λ-dynamics can simultaneously sample modifications at multiple sites within a chemical system. This enables base changes at different RNA sequence positions to yield free energy results for multiple modification combinations. There are limitations to λ-dynamics as well. Many of the calculated free energy differences, such as with uridine bound to the inosine antibody (Fig. 3A) or with $m^6{}_2A$ bound to the m^6A antibody (Fig. 4), predicted greater enhancement of binding than what was observed *in vitro* (Fig. 5). The starting models for the λ-dynamics calculations were based on the crystal structures of antibody fragments bound to nucleosides, but binding was tested *in vitro* with RNA oligos. This omission of the RNA phosphate backbone from the model, as well as the potential for sporadic self-associations or secondary structures in the unbound oligo, may have impacted the true binding values. The kinetics of RNA–protein complex assembly are also not modeled and may have unappreciated properties (e.g. [104]) that affect binding in the *in vitro* assays. Additional work probing RNA–protein interactions with λ-dynamics will undoubtedly improve these simulations. Moreover, the refinement of MD force fields [79, 76, 105, 106], particularly with respect to nucleic acids [79, 81, 106, 107], is a bustling area of research, and future advancements promise to further enhance the accuracy of these simulations.

While hundreds of RNA modifications have been identified, only a few dozen are available for experimental testing *in vitro*. Novel methods must be developed to examine how all modifications affect molecular interactions to decipher their biological mechanisms. This study establishes a workflow for using λ-dynamics to probe nucleic acid–protein interactions *in silico* (Fig. 6). The combinatorial efficiency of λ-dynamics enables rapid *in silico* examination of currently known and newly discovered RNA modifications. With high-resolution structures of nucleic acid–protein complexes, modified and unmodified nucleoside bases can be probed to explore how chemical changes to RNA affect protein-binding interactions. A library of RNA modifications can be tested at different RNA base positions to discover how these chemical changes may prevent or enhance protein binding. The development of accurate protein folding predictors, such as RoseTTAFold [108] and AlphaFold [109], now can provide accessible starting structural models for RNA–protein complexes to use with λ-dynamics. This computational approach can be used for DNA or RNA and is not limited by available chemistry. The work presented demonstrates how this strategy can probe for the specificity of antibodies. Future work can utilize this method to test how hundreds of RNA modifications affect their molecular interactions with any RNA-binding protein or other nucleic acids, delivering novel insights into their molecular functions.

Acknowledgements

The authors thank Dr. Millie Georgiadis for help with crystal data collection and model building, Dr. Giovanni Gonzalez-Gutierrez and the Physical Biochemistry Instrumentation Facility at Indiana University Bloomington for help with microscale thermophoresis, and members of the Aoki, Vilseck, and Zhang Labs for their helpful discussion. Data collection and processing was also facilitated by Craig A. Bingman of the Department of Biochemistry Collaborative Crystallography Core and the University of Wisconsin, which is supported by user fees and the department. X-ray diffraction data were collected on Advanced Photon Source beamline LRL-CAT 31-ID and the ESRF ID30B operated by the Life Sciences Collaborative Access Team (LS-CAT). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Company, which operates the facility. The authors acknowledge the Indiana University Pervasive Technology Institute for providing supercomputing and storage resources that have contributed to the research results reported within this paper.

Author contributions: M.A. carried out data curation, formal analyses, investigation, software, validation, visualization, writing the original draft, and writing revision and editing. W.Z. performed project conceptualization, resources, writing revision, and editing. J.V. executed project conceptualization, data curation, methodology, project administration, software, validation, visualization, and writing revision and editing. S.T.A. performed project conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validiation, writing original draft, and writing revision and editing.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

National Institute of General Medical Sciences (Grant/Award Number: 'R35GM142691,' 'R35GM146888'); Basic Energy Sciences (Grant/Award Number: 'DE-AC02-06CH11357'); School of Medicine, Indiana University (Grant/Award Number: 'Precision Health Initiative,' 'Research Support Funds Grant'). Funding to pay the Open Access publication charges for this article was provided by NIH/NIGMS (R35GM142691, R35GM146888) and Indiana University start up funds.

Data availability

The final structures and merged reflection files are deposited at wwPDB (wwpdb.org; PDB IDs: 8SIP, 8TCA, 8VEV). Unmerged reflection data were deposited at Integrated Resource for Reproducibility in Macromolecular Crystallography (proteindiffraction.org).

References

1. McCown PJ, Ruszkowska A, Kunkler CN *et al.* Naturally occurring modified ribonucleosides. *WIREs RNA* 2020;**11**:e1595. https://doi.org/10.1002/wrna.1595
2. Cappannini A, Ray A, Purta E *et al.* MODOMICS: a database of RNA modifications and related information. 2023 update.

Nucleic Acids Res 2024;**52**:D239–44. https://doi.org/10.1093/nar/gkad1083
3. Roundtree IA, Evans ME, Pan T *et al.* Dynamic RNA modifications in gene expression regulation. *Cell* 2017;**169**:1187–200. https://doi.org/10.1016/j.cell.2017.05.045
4. Kennedy EM, Courtney DG, Tsai K *et al.* Viral epitranscriptomics. *J Virol* 2017;**91**:e02263-16. https://doi.org/10.1128/JVI.02263-16
5. Wu X, Sang L, Gong Y. N6-methyladenine RNA modification and cancers. *Am J Cancer Res* 2018;**8**:1957–66.
6. Liu C, Sun H, Yi Y *et al.* Absolute quantification of single-base m(6)A methylation in the mammalian transcriptome using GLORI. *Nat Biotechnol* 2023;**41**:355–66. https://doi.org/10.1038/s41587-022-01487-9
7. Carlile TM, Rojas-Duran MF, Zinshteyn B *et al.* Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 2014;**515**:143–6. https://doi.org/10.1038/nature13802
8. Stoiber M, Quick J, Egan R *et al. de novo* identification of DNA modifications enabled by genome-guided nanopore signal processing. bioRxiv, https://doi.org/10.1101/094672, 10 April 2017, preprint: not peer reviewed.
9. Lorenz DA, Sathe S, Einstein JM *et al.* Direct RNA sequencing enables m(6)A detection in endogenous transcript isoforms at base-specific resolution. *RNA* 2020;**26**:19–28. https://doi.org/10.1261/rna.072785.119
10. Gao Y, Liu X, Wu B *et al.* Quantitative profiling of N(6)-methyladenosine at single-base resolution in stem-differentiating xylem of Populus trichocarpa using Nanopore direct RNA sequencing. *Genome Biol* 2021;**22**:22. https://doi.org/10.1186/s13059-020-02241-7
11. Hendra C, Pratanwanich PN, Wan YK *et al.* Detection of m6A from direct RNA sequencing using a multiple instance learning framework. *Nat Methods* 2022;**19**:1590–8. https://doi.org/10.1038/s41592-022-01666-1
12. Leger A, Amaral PP, Pandolfini L *et al.* RNA modifications detection by comparative nanopore direct RNA sequencing. *Nat Commun* 2021;**12**:7198. https://doi.org/10.1038/s41467-021-27393-3
13. Pratanwanich PN, Yao F, Chen Y *et al.* Identification of differential RNA modifications from nanopore direct RNA sequencing with xPore. *Nat Biotechnol* 2021;**39**:1394–402. https://doi.org/10.1038/s41587-021-00949-w
14. Liu H, Begik O, Lucas MC *et al.* Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nat Commun* 2019;**10**:4079. https://doi.org/10.1038/s41467-019-11713-9
15. Parker MT, Knop K, Sherwood AV *et al.* Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. *eLife* 2020;**9**:e49658. https://doi.org/10.7554/eLife.49658
16. Price AM, Hayer KE, McIntyre ABR *et al.* Direct RNA sequencing reveals m(6)A modifications on adenovirus RNA are necessary for efficient splicing. *Nat Commun* 2020;**11**:6016. https://doi.org/10.1038/s41467-020-19787-6
17. Jenjaroenpun P, Wongsurawat T, Wadley TD *et al.* Decoding the epitranscriptional landscape from native RNA sequences. *Nucleic Acids Res* 2021;**49**:e7. https://doi.org/10.1093/nar/gkaa620
18. Zhong ZD, Xie YY, Chen HX *et al.* Systematic comparison of tools used for m(6)A mapping from nanopore direct RNA sequencing. *Nat Commun* 2023;**14**:1906. https://doi.org/10.1038/s41467-023-37596-5
19. Sarkar A, Gasperi W, Begley U *et al.* Detecting the epitranscriptome. *WIREs RNA* 2021;**12**:e1663. https://doi.org/10.1002/wrna.1663
20. Horowitz S, Horowitz A, Nilsen TW *et al.* Mapping of N6-methyladenosine residues in bovine prolactin mRNA. *Proc Natl Acad Sci USA* 1984;**81**:5667–71. https://doi.org/10.1073/pnas.81.18.5667
21. Dominissini D, Moshitch-Moshkovitz S, Schwartz S *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 2012;**485**:201–6. https://doi.org/10.1038/nature11112
22. Grozhik AV, Olarerin-George AO, Sindelar M *et al.* Antibody cross-reactivity accounts for widespread appearance of m(1)A in 5′UTRs. *Nat Commun* 2019;**10**:5126. https://doi.org/10.1038/s41467-019-13146-w
23. Safra M, Sas-Chen A, Nir R *et al.* The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* 2017;**551**:251–5. https://doi.org/10.1038/nature24456
24. Li X, Xiong X, Wang K *et al.* Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat Chem Biol* 2016;**12**:311–6. https://doi.org/10.1038/nchembio.2040
25. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S *et al.* The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 2016;**530**:441–6. https://doi.org/10.1038/nature16998
26. Delatte B, Wang F, Ngoc LV *et al.* RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 2016;**351**:282–5. https://doi.org/10.1126/science.aac5253
27. Amort T, Rieder D, Wille A *et al.* Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol* 2017;**18**:1. https://doi.org/10.1186/s13059-016-1139-1
28. Stanfield RL, Wilson IA. Antibody structure. *Microbiol Spectr* 2014;**2**. https://doi.org/10.1128/microbiolspec.AID-0012-2013
29. Lipman NS, Jackson LR, Trudel LJ *et al.* Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources. *ILAR J* 2005;**46**:258–68. https://doi.org/10.1093/ilar.46.3.258
30. Lee CC, Perchiacca JM, Tessier PM. Toward aggregation-resistant antibodies by design. *Trends Biotechnol* 2013;**31**:612–20. https://doi.org/10.1016/j.tibtech.2013.07.002
31. Wilson IA, Stanfield RL. 50 Years of structural immunology. *J Biol Chem* 2021;**296**:100745. https://doi.org/10.1016/j.jbc.2021.100745
32. Pokkuluri PR, Bouthillier F, Li Y *et al.* Preparation, characterization and crystallization of an antibody Fab fragment that recognizes RNA. Crystal structures of native Fab and three Fab-mononucleotide complexes. *J Mol Biol* 1994;**243**:283–97. https://doi.org/10.1006/jmbi.1994.1654
33. Weichmann F, Hett R, Schepers A *et al.* Validation strategies for antibodies targeting modified ribonucleotides. *RNA* 2020;**26**:1489–506. https://doi.org/10.1261/rna.076026.120
34. Helm M, Motorin Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat Rev Genet* 2017;**18**:275–91. https://doi.org/10.1038/nrg.2016.169
35. Sonenberg N. eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem Cell Biol* 2008;**86**:178–83. https://doi.org/10.1139/O08-034
36. Song LF, Merz KM Jr. Evolution of alchemical free energy methods in drug discovery. *J Chem Inf Model* 2020;**60**:5308–18. https://doi.org/10.1021/acs.jcim.0c00547
37. Chodera JD, Mobley DL, Shirts MR *et al.* Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol* 2011;**21**:150–60. https://doi.org/10.1016/j.sbi.2011.01.011
38. Kollman P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem Rev* 1993;**93**:2395–417. https://doi.org/10.1021/cr00023a004
39. Knight JL, Brooks CL. Multisite λ dynamics for simulated structure–Activity relationship studies. *J Chem Theory Comput* 2011;**7**:2728–39. https://doi.org/10.1021/ct200444f
40. Knight JL, Brooks CL. Applying efficient implicit nongeometric constraints in alchemical free energy simulations. *J Comput Chem* 2011;**32**:3423–32. https://doi.org/10.1002/jcc.21921

41. Kong X, Brooks CL. λ-dynamics: a new approach to free energy calculations. *J Chem Phys* 1996;**105**:2414–23. https://doi.org/10.1063/1.472109
42. Knight JL, Brooks CL. λ-dynamics free energy simulation methods. *J Comput Chem* 2009;**30**:1692–700. https://doi.org/10.1002/jcc.21295
43. Vilseck JZ, Armacost KA, Hayes RL *et al.* Predicting binding free energies in a large combinatorial chemical space using multisite λ dynamics. *J Phys Chem Lett* 2018;**9**:3328–32. https://doi.org/10.1021/acs.jpclett.8b01284
44. Lee W-G, Gallardo-Macias R, Frey KM *et al.* Picomolar inhibitors of HIV reverse transcriptase featuring bicyclic replacement of a cyanovinylphenyl group. *J Am Chem Soc* 2013;**135**:16705–13. https://doi.org/10.1021/ja408917n
45. Lee W-G, Frey KM, Gallardo-Macias R *et al.* Picomolar inhibitors of HIV-1 reverse transcriptase: design and crystallography of naphthyl phenyl ethers. *ACS Med Chem Lett* 2014;**5**:1259–62. https://doi.org/10.1021/ml5003713
46. Vilseck JZ, Sohail N, Hayes RL *et al.* Overcoming challenging substituent perturbations with multisite λ-dynamics: a case study targeting β-secretase 1. *J Phys Chem Lett* 2019;**10**:4875–80. https://doi.org/10.1021/acs.jpclett.9b02004
47. Keränen H, Pérez-Benito L, Ciordia M *et al.* Acylguanidine beta secretase 1 inhibitors: a combined experimental and free energy perturbation study. *J Chem Theory Comput* 2017;**13**:1439–53. https://doi.org/10.1021/acs.jctc.6b01141
48. Hanquier JN, Sanders K, Berryhill CA *et al.* Identification of nonhistone substrates of the lysine methyltransferase PRDM9. *J Biol Chem* 2023;**299**:104651. https://doi.org/10.1016/j.jbc.2023.104651
49. Peck Justice SA, Barron MP, Qi GD *et al.* Mutant thermal proteome profiling for characterization of missense protein variants and their associated phenotypes within the proteome. *J Biol Chem* 2020;**295**:16219–38. https://doi.org/10.1074/jbc.RA120.014576
50. Hayes RL, Vilseck JZ, Brooks CL 3rd. Approaching protein design with multisite lambda dynamics: accurate and scalable mutational folding free energies in T4 lysozyme. *Protein Sci* 2018;**27**:1910–22. https://doi.org/10.1002/pro.3500
51. Bandeira N, Pham V, Pevzner P *et al.* Automated de novo protein sequencing of monoclonal antibodies. *Nat Biotechnol* 2008;**26**:1336–8. https://doi.org/10.1038/nbt1208-1336
52. Castellana NE, Pham V, Arnott D *et al.* Template proteogenomics: sequencing whole proteins using an imperfect database. *Mol Cell Proteomics* 2010;**9**:1260–70. https://doi.org/10.1074/mcp.M900504-MCP200
53. Castellana NE, McCutcheon K, Pham VC *et al.* Resurrection of a clinical antibody: template proteogenomic de novo proteomic sequencing and reverse engineering of an anti-lymphotoxin-alpha antibody. *Proteomics* 2011;**11**:395–405. https://doi.org/10.1002/pmic.201000487
54. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:125–32. https://doi.org/10.1107/S0907444909047337
55. Emsley P, Lohkamp B, Scott WG *et al.* Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:486–501. https://doi.org/10.1107/S0907444910007493
56. Liebschner D, Afonine PV, Baker ML *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 2019;**75**:861–77. https://doi.org/10.1107/S2059798319011471
57. Williams CJ, Headd JJ, Moriarty NW *et al.* MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci* 2018;**27**:293–315. https://doi.org/10.1002/pro.3330
58. Olsson MHM, Søndergaard CR, Rostkowski M *et al.* PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput* 2011;**7**:525–37. https://doi.org/10.1021/ct100578z
59. Sondergaard CR, Olsson MH, Rostkowski M *et al.* Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J Chem Theory Comput* 2011;**7**:2284–95. https://doi.org/10.1021/ct200133y
60. Jo S, Kim T, Iyer VG *et al.* CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 2008;**29**:1859–65. https://doi.org/10.1002/jcc.20945
61. Brooks BR, Brooks CL 3rd, Mackerell AD Jr *et al.* CHARMM: the biomolecular simulation program. *J Comput Chem* 2009;**30**:1545–614. https://doi.org/10.1002/jcc.21287
62. Brooks BR, Bruccoleri RE, Olafson BD *et al.* CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 1983;**4**:187–217. https://doi.org/10.1002/jcc.540040211
63. Hayes RL, Buckner J, Brooks CL 3rd. BLaDE: a basic Lambda Dynamics engine for GPU-accelerated molecular dynamics free energy calculations. *J Chem Theory Comput* 2021;**17**:6799–807. https://doi.org/10.1021/acs.jctc.1c00833
64. Åqvist J, Wennerström P, Nervall M *et al.* Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem Phys Lett* 2004;**384**:288–94. https://doi.org/10.1016/j.cplett.2003.12.039
65. Chow K-H, Ferguson DM. Isothermal-isobaric molecular dynamics simulations with Monte Carlo volume sampling. *Comput Phys Commun* 1995;**91**:283–9. https://doi.org/10.1016/0010-4655(95)00059-O
66. Leimkuhler B, Matthews C. Efficient molecular dynamics using geodesic integration and solvent-solute splitting. *Proc Math Phys Eng Sci* 2016;**472**:20160138.
67. Ryckaert J-P, Ciccotti G, Berendsen HJ. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 1977;**23**:327–41. https://doi.org/10.1016/0021-9991(77)90098-5
68. van Gunsteren WF, Berendsen HJ. Algorithms for macromolecular dynamics and constraint dynamics. *Mol Phys* 1977;**34**:1311–27. https://doi.org/10.1080/00268977700102571
69. Andersen HC. Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. *J Comput Phys* 1983;**52**:24–34. https://doi.org/10.1016/0021-9991(83)90014-1
70. Kräutler V, Van Gunsteren WF, Hünenberger PH. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem* 2001;**22**:501–8. https://doi.org/10.1002/1096-987X(20010415)22:5(501::AID-JCC1021)3.0.CO;2-V
71. Darden T, York D, Pedersen L. Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems. *J Chem Phys* 1993;**98**:10089–92. https://doi.org/10.1063/1.464397
72. Essmann U, Perera L, Berkowitz ML *et al.* A smooth particle mesh Ewald method. *J Chem Phys* 1995;**103**:8577–93. https://doi.org/10.1063/1.470117
73. Huang Y, Chen W, Wallace JA *et al.* All-atom continuous constant pH molecular dynamics with particle mesh ewald and titratable water. *J Chem Theory Comput* 2016;**12**:5411–21. https://doi.org/10.1021/acs.jctc.6b00552
74. Steinbach PJ, Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem* 1994;**15**:667–83. https://doi.org/10.1002/jcc.540150702
75. Jorgensen W, Chandrasekhar J, Madura J *et al.* Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;**79**:926–35. https://doi.org/10.1063/1.445869
76. Huang J, Rauscher S, Nawrocki G *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 2017;**14**:71–3. https://doi.org/10.1038/nmeth.4067
77. Mackerell AD Jr, Feig M, Brooks CL. Iii Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein

conformational distributions in molecular dynamics simulations. *J Comput Chem* 2004;**25**:1400–15. https://doi.org/10.1002/jcc.20065
78. Best RB, Zhu X, Shim J *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1 and χ2 dihedral angles. *J Chem Theory Comput* 2012;**8**:3257–73. https://doi.org/10.1021/ct300400x
79. Denning EJ, Priyakumar UD, Nilsson L *et al.* Impact of 2′-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* 2011;**32**:1929–43. https://doi.org/10.1002/jcc.21777
80. Foloppe N, MacKerell JAD. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 2000;**21**:86–104. https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2%3c86::AID-JCC2%3e3.0.CO;2-G
81. Xu Y, Vanommeslaeghe K, Aleksandrov A *et al.* Additive CHARMM force field for naturally occurring modified ribonucleotides. *J Comput Chem* 2016;**37**:896–912. https://doi.org/10.1002/jcc.24307
82. Vilseck JZ, Cervantes LF, Hayes RL *et al.* Optimizing multisite λ-dynamics throughput with charge renormalization. *J Chem Inf Model* 2022;**62**:1479–88. https://doi.org/10.1021/acs.jcim.2c00047
83. Hayes RL, Brooks CL 3rd. A strategy for proline and glycine mutations to proteins with alchemical free energy calculations. *J Comput Chem* 2021;**42**:1088–94. https://doi.org/10.1002/jcc.26525
84. Rocklin GJ, Mobley DL, Dill KA *et al.* Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. *J Chem Phys* 2013;**139**:184103. https://doi.org/10.1063/1.4826261
85. Chen W, Deng Y, Russell E *et al.* Accurate calculation of relative binding free energies between ligands with different net charges. *J Chem Theory Comput* 2018;**14**:6346–58. https://doi.org/10.1021/acs.jctc.8b00825
86. Petrov D, Perthold JW, Oostenbrink C *et al.* Guidelines for free-energy calculations involving charge changes. *J Chem Theory Comput* 2024;**20**:914–25. https://doi.org/10.1021/acs.jctc.3c00757
87. Hayes RL, Armacost KA, Vilseck JZ *et al.* Adaptive landscape flattening accelerates sampling of alchemical space in multisite lambda dynamics. *J Phys Chem B* 2017;**121**:3626–35. https://doi.org/10.1021/acs.jpcb.6b09656
88. Kumar S, Rosenberg JM, Bouzida D *et al.* THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 1992;**13**:1011–21. https://doi.org/10.1002/jcc.540130812
89. Fanning SW, Horn JR. An anti-hapten camelid antibody reveals a cryptic binding site with significant energetic contributions from a nonhypervariable loop. *Protein Sci* 2011;**20**:1196–207. https://doi.org/10.1002/pro.648
90. Yau KY, Lee H, Hall JC. Emerging trends in the synthesis and improvement of hapten-specific recombinant antibodies. *Biotechnol Adv* 2003;**21**:599–637. https://doi.org/10.1016/S0734-9750(03)00104-6
91. Patil DP, Pickering BF, Jaffrey SR. Reading m(6)A in the transcriptome: m(6)A-binding proteins. *Trends Cell Biol* 2018;**28**:113–27. https://doi.org/10.1016/j.tcb.2017.10.001
92. Li F, Zhao D, Wu J *et al.* Structure of the YTH domain of human YTHDF2 in complex with an m(6)A mononucleotide reveals an aromatic cage for m(6)A recognition. *Cell Res* 2014;**24**:1490–2. https://doi.org/10.1038/cr.2014.153
93. Luo S, Tong L. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc Natl Acad Sci USA* 2014;**111**:13834–9. https://doi.org/10.1073/pnas.1412742111
94. Xu C, Wang X, Liu K *et al.* Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat Chem Biol* 2014;**10**:927–9. https://doi.org/10.1038/nchembio.1654
95. Nicastro G, Abis G, Klein P *et al.* Direct m6A recognition by IMP1 underlays an alternative model of target selection for non-canonical methyl-readers. *Nucleic Acids Res* 2023;**51**:8774–86. https://doi.org/10.1093/nar/gkad534
96. Xu C, Liu K, Ahmed H *et al.* Structural basis for the discriminative recognition of N6-methyladenosine RNA by the Human YT521-B homology Domain Family of proteins. *J Biol Chem* 2015;**290**:24902–13. https://doi.org/10.1074/jbc.M115.680389
97. Wienken CJ, Baaske P, Rothbauer U *et al.* Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* 2010;**1**:100. https://doi.org/10.1038/ncomms1093
98. Seeliger D, Buelens FP, Goette M *et al.* Towards computational specificity screening of DNA-binding proteins. *Nucleic Acids Res* 2011;**39**:8281–90. https://doi.org/10.1093/nar/gkr531
99. Kappel K, Jarmoskaite I, Vaidyanathan PP *et al.* Blind tests of RNA-protein binding affinity prediction. *Proc Natl Acad Sci USA* 2019;**116**:8336–41. https://doi.org/10.1073/pnas.1819047116
100. Hajnic M, Alonso-Gil S, Polyansky AA *et al.* Interaction preferences between protein side chains and key epigenetic modifications 5-methylcytosine, 5-hydroxymethycytosine and N(6)-methyladenine. *Sci Rep* 2022;**12**:19583. https://doi.org/10.1038/s41598-022-23585-z
101. Beierlein FR, Kneale GG, Clark T. Predicting the effects of basepair mutations in DNA-protein complexes by thermodynamic integration. *Biophys J* 2011;**101**:1130–8. https://doi.org/10.1016/j.bpj.2011.07.003
102. Gapsys V, de Groot BL. Alchemical free energy calculations for nucleotide mutations in protein-DNA complexes. *J Chem Theory Comput* 2017;**13**:6275–89. https://doi.org/10.1021/acs.jctc.7b00849
103. Olson MA. Calculations of free-energy contributions to protein-RNA complex stabilization. *Biophys J* 2001;**81**:1841–53. https://doi.org/10.1016/S0006-3495(01)75836-5
104. James LC, Roversi P, Tawfik DS. Antibody multispecificity mediated by conformational diversity. *Science* 2003;**299**:1362–7. https://doi.org/10.1126/science.1079731
105. Maier JA, Martinez C, Kasavajhala K *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* 2015;**11**:3696–713. https://doi.org/10.1021/acs.jctc.5b00255
106. Zgarbova M, Otyepka M, Sponer J *et al.* Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* 2011;**7**:2886–902. https://doi.org/10.1021/ct200162x
107. Aduri R, Psciuk BT, Saro P *et al.* AMBER force field parameters for the naturally occurring modified nucleosides in RNA. *J Chem Theory Comput* 2007;**3**:1464–75. https://doi.org/10.1021/ct600329w
108. Baek M, DiMaio F, Anishchenko I *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021;**373**:871–6. https://doi.org/10.1126/science.abj8754
109. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500. https://doi.org/10.1038/s41586-024-07487-w

Received: January 27, 2024. **Revised:** February 19, 2025. **Editorial Decision:** February 19, 2025. **Accepted:** February 20, 2025