

Visualizing RNA conformational and architectural heterogeneity in solution

Jienyu Ding

NCI

Yun-Tzai Lee

NIH NCI Frederick <https://orcid.org/0000-0001-8123-6207>

Y Bhandari

NCI

Lixin Fan

Basic Science Program, Frederick National Laboratory for Cancer Research, SAXS Core Facility, Center for Cancer Research of the National Cancer Institute, Frederick, MD

Charles Schwieters

National Institutes of Health <https://orcid.org/0000-0002-4216-4658>

Ping Yu

NCI, NIH

Sergey Tarasov

NCI-Frederick

Jason Stagno

National Cancer Institute, NIH <https://orcid.org/0000-0002-6464-7829>

Buyong Ma

Shanghai Jiaotong University <https://orcid.org/0000-0002-7383-719X>

Ruth Nussinov

Frederick National Laboratory for Cancer Research

Alan Rein

<https://orcid.org/0000-0002-8273-546X>

Jinwei Zhang

National Institutes of Health <https://orcid.org/0000-0002-2114-173X>

Yun-Xing Wang (✉ wangyunx@mail.nih.gov)

NCI, NIH <https://orcid.org/0000-0002-2175-0148>

Article

Keywords:

Posted Date: April 8th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1506923/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Visualizing RNA conformational and architectural heterogeneity in solution

Jienyu Ding¹, Yun-Tzai Lee¹, Yuba Bhandari¹, Lixin Fan³, Charles D. Schwieters², Ping Yu¹, Sergey G. Tarosov⁴, Jason R. Stagno¹, Buyong Ma⁵, Ruth Nussinov⁵, Alan Rein⁶, Jinwei Zhang⁷, Yun-Xing Wang¹

¹ Protein-Nucleic Acid Interaction Section, Center for Structural Biology, National Cancer Institute, Frederick, Maryland 21702, USA; ² Computational Biomolecular Magnetic Resonance Core, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA; ³ Basic Science Program, Frederick National Laboratory for Cancer Research, Small Angle X-ray Scattering Core Facility of National Cancer Institute , National Institutes of Health, Frederick, Maryland 21702, USA ⁴Biophysical Resource, Center for Structural Biology National Cancer Institute, Frederick, Maryland 21702, USA;

⁵Computational Structural Biology Section, Frederick National Laboratory for Cancer Research in the Laboratory of Cancer Immunometabolism, National Cancer Institute, Frederick, MD 21702, USA; ⁶Retrovirus Assembly Section, HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702, USA; ⁷Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA

* Yun-Xing Wang, wangyunx@mail.nih.gov

Abstract

RNA flexibility is reflected in its heterogeneous conformation. Existing techniques, such as NMR, crystallography, or cryo-EM, are not suited for studying highly heterogeneous conformation. Through direct visualization using atomic force microscopy (AFM) and the adenosylcobalamin riboswitch aptamer domain as an example, we show that a single RNA sequence folds into conformationally and architecturally heterogeneous structures under near-physiological solution conditions. Recapitulated 3D topological structures from AFM molecular surfaces reveal that all conformers share the same secondary structural elements. Only a population-weighted cohort, not any single conformer, including the crystal structure, can account for the ensemble behaviors. All conformers, except for one, are functionally active in terms of ligand binding. Our findings provide direct visual evidence that the sequence-structure relationship of RNA under physiologically relevant solution conditions is more complex than the one-to-one relationship for well-structured proteins. The direct visualization of conformational and architectural ensembles in solution may prompt a revision to our current views and approaches to RNA structural analyses.

Main

Ever since the first protein structure was determined in 1958¹, we have become accustomed to the idea that a given primary amino acid sequence folds into corresponding secondary structures and a cognate tertiary structure of a well-folded protein. This idea has been reinforced by the recent, rapid, and exciting developments in predicting protein structures from primary sequences and extensive knowledge available in structural databases²⁻⁴. For RNA, the functional roles of structural dynamics continue to emerge in a number of biological and cellular contexts³. RNA differs from protein in several fundamental aspects that dictate its folding. It is a negatively charged polyelectrolyte, highly hydrophilic, and made of only four different building blocks that result in a highly degenerate folding-energy landscape. That is, one RNA primary sequence may adopt more than one secondary, tertiary, or higher-order structure, whose architectures, populations of various conformers, and motion timescales have all been shown to contribute to the versatile functions of RNA. Indeed, extensive and pioneering studies have revealed that RNA structures are far more dynamic and heterogeneous than what could be described by single conformers⁵⁻⁸. Nevertheless, those studies provide only time-averaged behaviors of ensembles, not explicit spatial descriptions about individual molecules or subgroups of conformers in low populations. Moreover, current high-resolution techniques for structure determination (NMR, crystallography, cryo-EM) rely on signal enhancements over a large number of molecules of the same or very similar conformation to improve the signal-to-noise ratio. These methods are limited to studies of relatively homogeneous samples either by driving the molecules to uniformity under extreme conditions or by removing heterogeneous species via purification. As a result, both the extent and nature of RNA structural heterogeneity under near-physiological conditions have largely remained unknown.

The question, therefore, of whether the description “one sequence, one structure” is appropriate for RNA, remains to be answered and becomes even more urgent given the major development in predicting RNA structures⁹. This question is fundamental to the understanding of functional RNA structural dynamics, and the answer requires direct examination of individual molecules in solution. The recent evolution of atomic force microscopy (AFM) enables such direct visualization with several clear advantages. Measurements can be performed in physiologically relevant buffer conditions; sample consumption is extremely low, requiring only microliter volumes at nanomolar concentrations; and molecules can be observed in their native states without any manipulation, e.g. labeling, freezing, staining, or crystallization¹⁰. Moreover, as data are recorded in real space, a high signal-to-noise level can be achieved from a single image. Thus, the solution AFM method is well suited for studying highly heterogeneous molecular systems under near-native conditions.

Cobalamin riboswitches are one of the most widely distributed riboswitches, regulating B12 biosynthesis in bacteria¹¹⁻¹⁴. Chemical probing, analytical size-exclusion chromatography and multi-angle light scattering (SEC-MALS) data showed that possible heterogeneous conformations for the RNA exist in solution¹⁵⁻¹⁷. In crystals, the adenosylcobalamin riboswitch aptamer domain from *Thermoanaerobacter tengcongensis* (rCbl), in complex with adenosylcobalamin (AdoCbl), exhibited a single and monomeric compact tertiary fold stabilized by a long-range intramolecular kissing-loop (KL) interaction¹⁸, whereas the structure of an atypical cobalamin riboswitch showed homodimeric intermolecular KL interactions¹⁷. Using the rCbl RNA as a proof-of-principle, our solution AFM data show that, in contrast to its protein

counterpart, a single RNA sequence can form multiple conformations and individually definable multimeric architectures under physiologically relevant Mg²⁺ concentration.

Results

Visualizing RNA conformational and architectural heterogeneity

High-resolution AFM images of the rCbl RNA reveal structural heterogeneity among thousands of particles, which are individually classified as monomeric (*Y-shape* (*Y*), *P-shape* (*P*), *candy-shape* (*candy* and *candy2*), *compact*), dimeric, or multimeric (**Fig. 1**, **Supplementary Fig. 1**). In the absence of ligand (2073 particles), the aptamer is almost exclusively monomeric, whereas, in the presence of ligand (3410 particles), it is predominantly dimeric, suggesting that the dimers are stabilized by ligand binding. Second, the presence of ligand reveals a marked decrease in the populations of *candy* (from 49.4 to 2.3%) and *P* (from 22.5 to 9.1%) monomer conformations, coupled with an increase (from 1.6 to 62.7%) in the dimer population, indicating that those monomer conformations may undergo ligand-induced structural changes that ultimately drive dimer formation. The third major species observed, *Y*, exhibits a very similar population (~22%) under both conditions, suggesting that its structure is independent of ligand and may represent a binding-incompetent or misfolded conformation. It is noteworthy to mention that, with direct visualization by AFM, minor species with only a single or few copies can be detected (**Fig. 1a**, white circles), and all conformers are present, albeit with different populations, under all conditions tested in this study with an extensive investigation of effects on the conformational heterogeneity by temperature, salt concentrations, folding processes, purification procedures and storage conditions (**Supplementary Fig. 2**).

Topological structures of conformers

We further examined the various conformers by recapitulating their 3D topological structures, one for each monomer and dimer class, using the coarse-grained dynamic fitting, in which the high-resolution single-particle AFM images were applied as topological constraints. Each class was determined to have a distinct structural fold (**Fig. 2, Supplementary Table 1**), while they all share the same secondary structure. The *compact* monomer with intramolecular KL interaction between L5 and L13 most closely resembles the ligand-bound crystal structure (PDB: 4GMA)¹⁸, with a cross-correlation score between the AFM image and the crystal structure of ~0.87. Interestingly, however, this crystal-structure-like monomer is present at an extremely low population in both the presence and absence of ligand at 1 mM Mg²⁺ concentration. All other monomer classes reveal more extended conformations, where L13 is disengaged from L5 (**Fig. 2**). In the cases of *P* and *candy*, the P13 helix becomes an extension of P1. As a result, L5 and L13 are free to form intermolecular KL interactions, as observed in the majority of dimer classes (**Figs. 2, 3**).

Heterogeneous conformers account for ensemble behaviors

The presence of multiple conformations in solution under physiological Mg²⁺ concentration is corroborated by SAXS measurements using the same batch of RNA. The topological AFM-derived structures and the volume fraction of each monomer and dimer conformation, in the absence or presence of ligand, were used to synthesize individual SAXS curves, which were then compared to the experimental SAXS data (**Fig. 2c, d**). None of the back-calculated SAXS curves of single species, including that of the crystal structure, could be fit to the experimental data, consistent with highly heterogeneous RNA conformations present in solution revealed by AFM.

The experimental data in both cases could only be approximated using a fractional combination of all monomer and dimer topological structures (**Fig. 2c,d; Supplementary Fig. 3**).

Heterogeneous conformers are active

The raw isotherm and thermogram from isothermal titration calorimetry (ITC) data (**Fig. 2e**) indicate that the reaction between aptamer and ligand does not follow a simple 1:1 binding regime but rather a mixture of endothermic and exothermic events. Singular value decomposition (SVD) analysis reveals that the data are adequately fit with a minimum of two principal components representing interconverting active conformations with different populations (**Fig. 2e, Supplementary Table 2**). Moreover, the resulting apparent binding stoichiometry (**Supplementary Table 2**, $N_{app}=0.78$) reflects that majority of heterogeneous species are active in terms of ligand binding, and only ~22% of the RNA is incapable of ligand binding, which is consistent with the *Y* conformer population observed by AFM that did not respond to ligand (**Fig. 1e, Supplementary Fig. 1a**). Of the binding-competent aptamers, the majority (component 1: ~69%) exhibit a typical sigmoidal isotherm signature of exothermic binding interaction with ligand, which likely involves only minor conformational changes. Component 2 (~19%) reflects an endothermic reaction, most probably associated with large conformational changes by an induced-fit binding scenario, followed by an exothermic reaction that is similar to component 1. The endothermic nature of this process is consistent with the enthalpic cost required to break existing contacts that stabilize the non-native conformations. The ITC data are consistent with the presence of heterogeneous conformational species and the ligand-binding-incompetent conformer.

Conserved RNA-RNA interactions drive the formation of architectural conformers

The promiscuity found in the dimer/multimer formations of rCbl is not caused by non-specific interactions, such as random base-pairing or electrostatic interactions, which one would otherwise suspect if no direct visual evidence were available. Instead, it is driven by well-known RNA-RNA interaction motifs. Most notable in this case is the KL interaction between L5 and L13 of two molecules, which appears in all but one of the dimer classes observed by AFM (**Figs. 2b and 3**). Specifically, this interaction drives the formation of the fully-formed (FF), one-leg-out (OLO) and hook (HK) dimers between two P-shape or two candy-shape homo-conformers, and P-Y and P-candy dimers between two hetero-conformers (**Fig. 3**). The OLO dimer is likely an FF dimer in the process of either forming or dissociating. Another well-known interaction motif that plays a role in rCbl dimer formation is the tetraloop-minor-groove interaction, as observed in the SS dimer (**Fig. 3**). Common to all these interactions is the requirement of chemical and structural complementarity between interacting partners. In essence, the RNA-RNA interactions are governed by principles similar to protein-protein interactions. These principles include complementarity of size, shape, chemical composition and structure^{19,20}. The interactions in RNA and proteins differ only in how specificity is achieved. In rCbl RNA, the specificity is achieved through both the chemical and structural complementarity such as the KL and the tetraloop-minor-groove interactions between structural elements, whereas in proteins they are largely driven by hydrophobic interfaces with size/shape complementarity.

Ligand-binding-incompetent Y-shape conformers

The topological structures of *Y* conformers, subclassified as *Y1* and *Y2* (**Fig. 4b**), both exhibit extended and partially folded structures lacking the L5/L13 KL interaction, yet with an intact P6 extension. The P6 extension contributes to one leg of the *Y*, while the other two legs could be either P2, P13, or P1 together with P13 (**Supplementary Fig. 4**). To further investigate the correlation between *Y* conformers and inactive species, we repeated AFM and ITC analyses for two mutants of rCbl, in which the essential KL interaction is abrogated by deleting P13 (M2) or by replacing L13 with a tetra-loop (M3) (**Supplementary Fig. 5**). As expected, AFM observation of either of these mutants shows >80% *Y* population, with the remainder of particles resembling *P* and *candy*, each at populations <10% without dimers (**Fig. 4c, Supplementary Fig. 6**). M2 *Y* conformers exhibit a larger mean particle volume, even though both mutant *Y* conformers have similar particle diameters (**Supplementary Fig. 7**), suggesting that the P13 deletion results in more extended conformations. This is further supported by the topological structures calculated from representative *Y* particles of each mutant (**Fig. 4b**), which show very similar conformations (also similar to rCbl), where the P6 extension constitutes one leg, and the other two legs are comprised of P2 and P4/P5 (M2) or P2 and P13 (M3), respectively (**Supplementary Fig. 4**). More importantly, as in the case of rCbl, the population of mutant *Y* conformers does not change in the presence of ligand (**Fig. 4c**), once again indicating that these *Y* conformers are a ligand-binding-incompetent species. This is further confirmed by ITC results where M3 shows no AdoCbl binding activity (**Fig. 4d**). The presence of ligand-binding-incompetent conformers was also reported in a close cousin of rCbl, env8 HyCbl¹⁵.

Discussion

Both RNA and protein can fold into three-dimensional (3D) structures. One of the basic premises of the revolutionary protein-structure prediction²⁻⁴ is that a primary sequence folds into a specific 3D structure under a physiological environment. Such a one-to-one relationship is appropriate for folded proteins, as clearly evidenced in the structure database. Although these advances in protein-structure prediction provide hope and encouragement for overcoming unique challenges in elucidating RNA 3D structures⁹, abundant experimental evidence suggests that a given RNA sequence can adopt diverse structural folds under physiologically relevant conditions⁵⁻⁸. Many of these functionally important folds might have been precluded from structure determination using current methods that require homogeneous samples. The direct visualization presented in our study illustrates heterogeneous conformation and architecture of a single RNA sequence in a near physiologically relevant solution condition. Our observation may prompt a revision to our current views and approaches to RNA structural analyses.

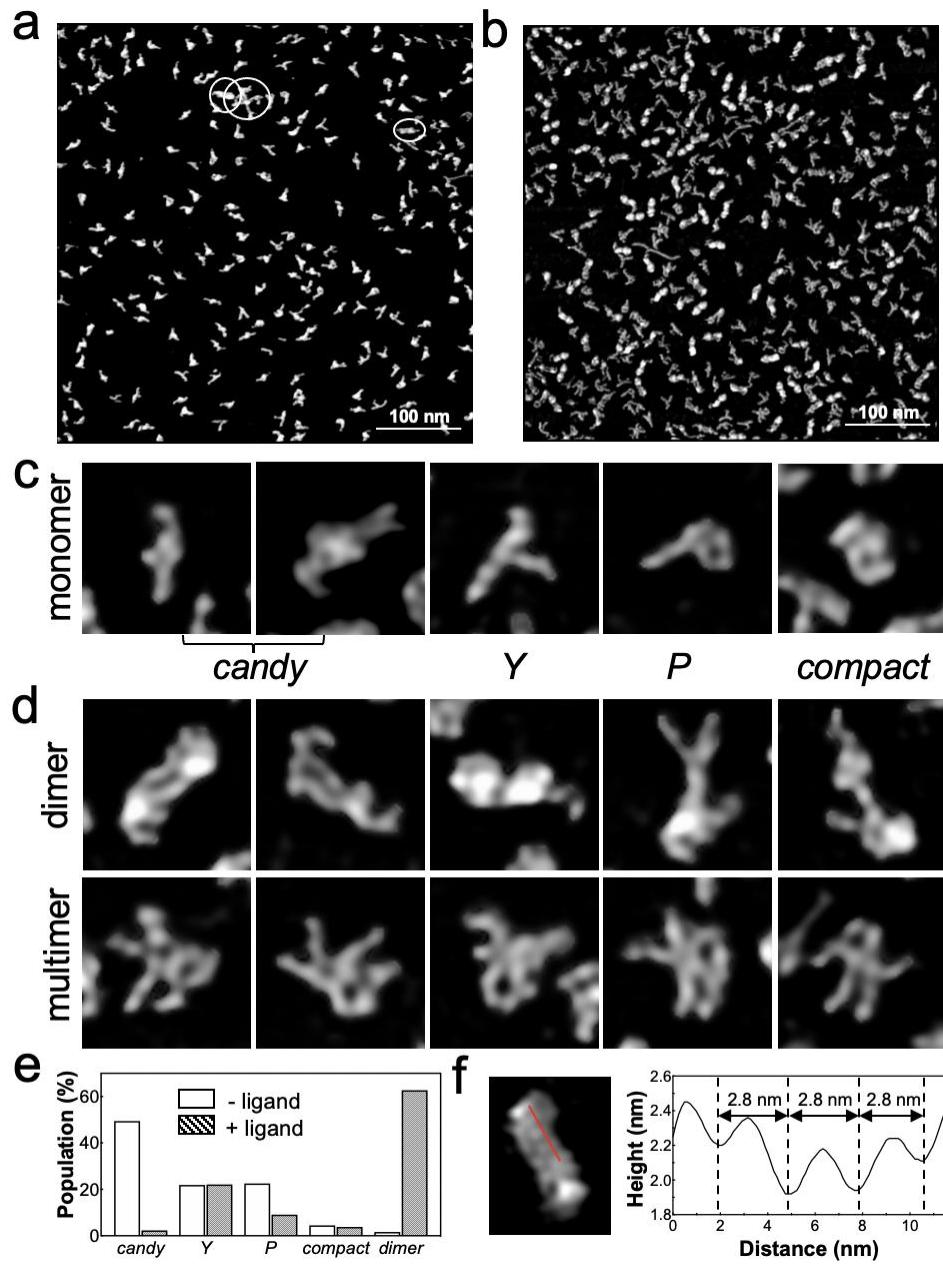


Figure 1 | Conformational heterogeneity of rCbl in solution. Sections of high-resolution AFM images of rCbl in the absence (**a**) or presence (**b**) of ligand. White circles indicate the detection of minor species down to only one or a few copies in solution. (**c**) Monomeric conformers. (**d**) Dimers classified based on shape. (**e**) Tallies of individual monomeric conformations (*candy*, *Y*, *P*, *compact*) and all dimeric species, in the absence (2073 particles) or presence (3410 particles) of ligand. (**f**) Height and distance parameters measured across the surface (red line) of a particle showing periodic corrugation of 2.8 nm, corresponding to RNA helical pitch.

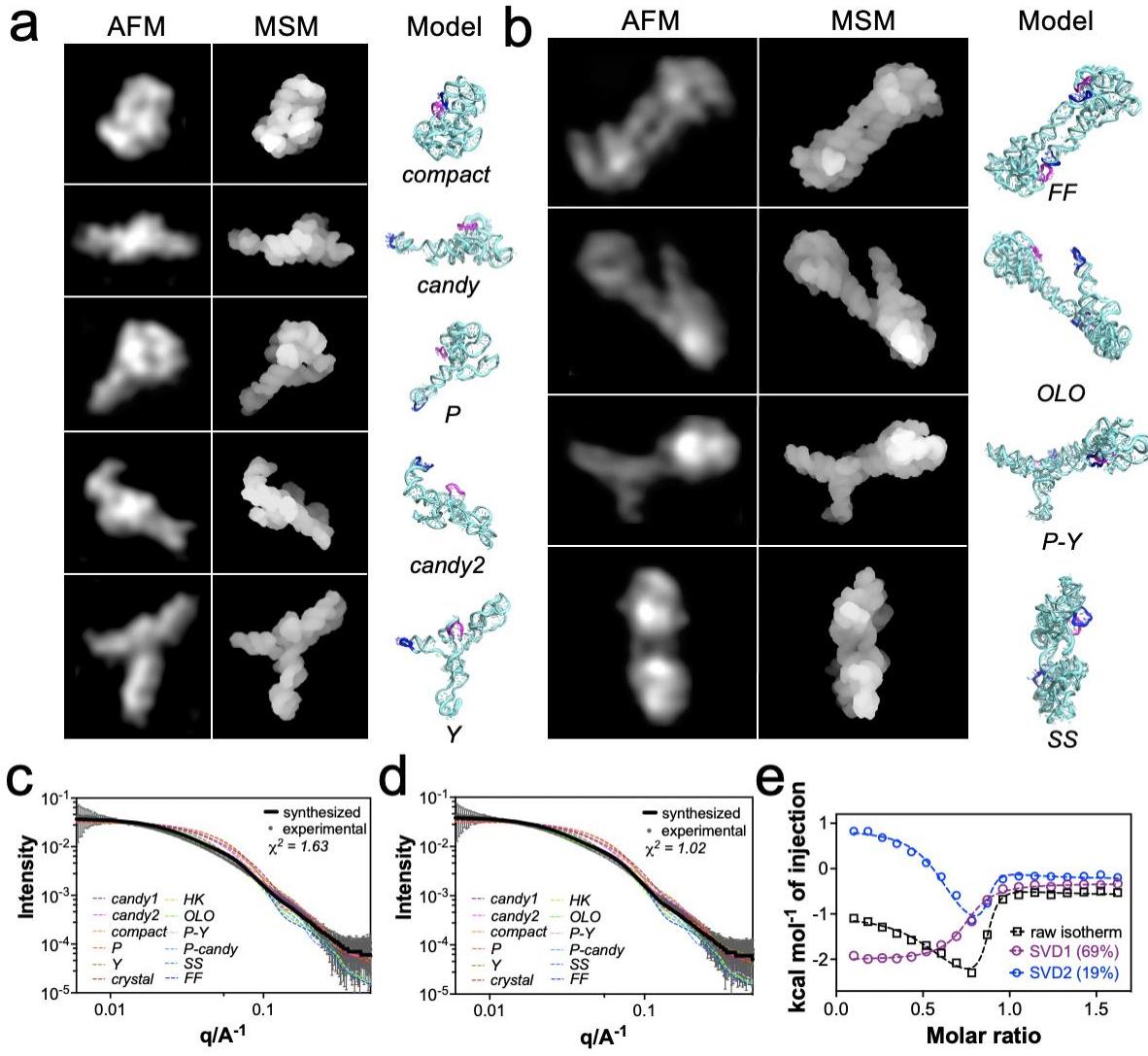


Figure 2 | Recapitulated 3D topological structures of the rCbl heterogeneous conformers corroborated by SAXS and ITC. (a-b) Experimental AFM images, the recapitulated structures in molecular surface model (MSM) and cartoon models for the monomeric (a) and dimeric (b) conformers. L5 and L13 are colored in magenta and blue, respectively. (c-d) Comparing the experimental SAXS curve (gray solid line) with back-calculated SAXS curves of the individual conformers (colored dash lines) with the synthesized SAXS curve (red solid line). Experimental SAXS data were recorded for rCbl (1 μM) in absence (c) or presence (d) of 10 μM AdoCbl. The synthesized SAXS curves were calculated using the structures of all conformers (**Methods**) with $\chi^2 = 1.68$ for rCbl in absence of ligand and $\chi^2 = 1.00$ for rCbl in presence of ligand. (e) Raw and deconvoluted (SVD1 and SVD2) isotherms of rCbl titrated with AdoCbl (**Supplementary Table 2**). The eigenvalue fractions for principal component1 (SVD1) and component2 (SVD2) are about 69% and 19%, respectively.

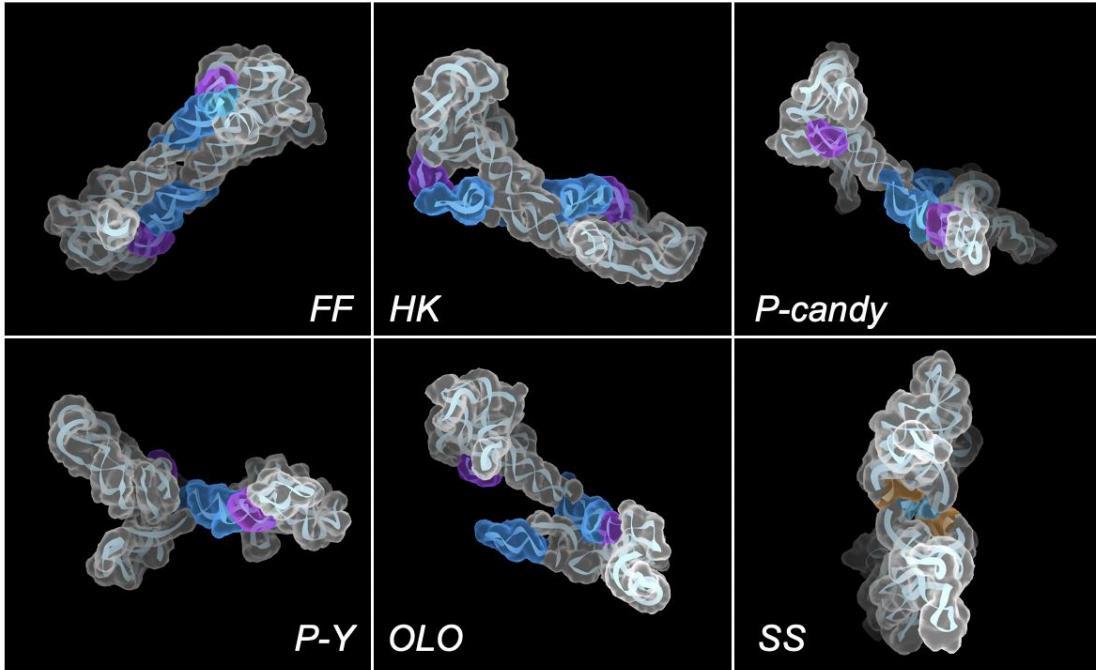


Figure 3 | Formation of the rCbl heterogeneous dimers through conserved RNA-RNA interaction motifs. Five of the six dimer classifications observed by AFM exhibit one or two KL interactions between L5/P5 (purple) and L13/P13 (blue): fully-formed (*FF*), hook (*HK*), *P-candy*, *P-Y*, and one-leg-out (*OLO*). The shoulder-to-shoulder (*SS*) dimer forms through symmetrical tetra-loop (cyan) - minor-groove (orange) interactions.

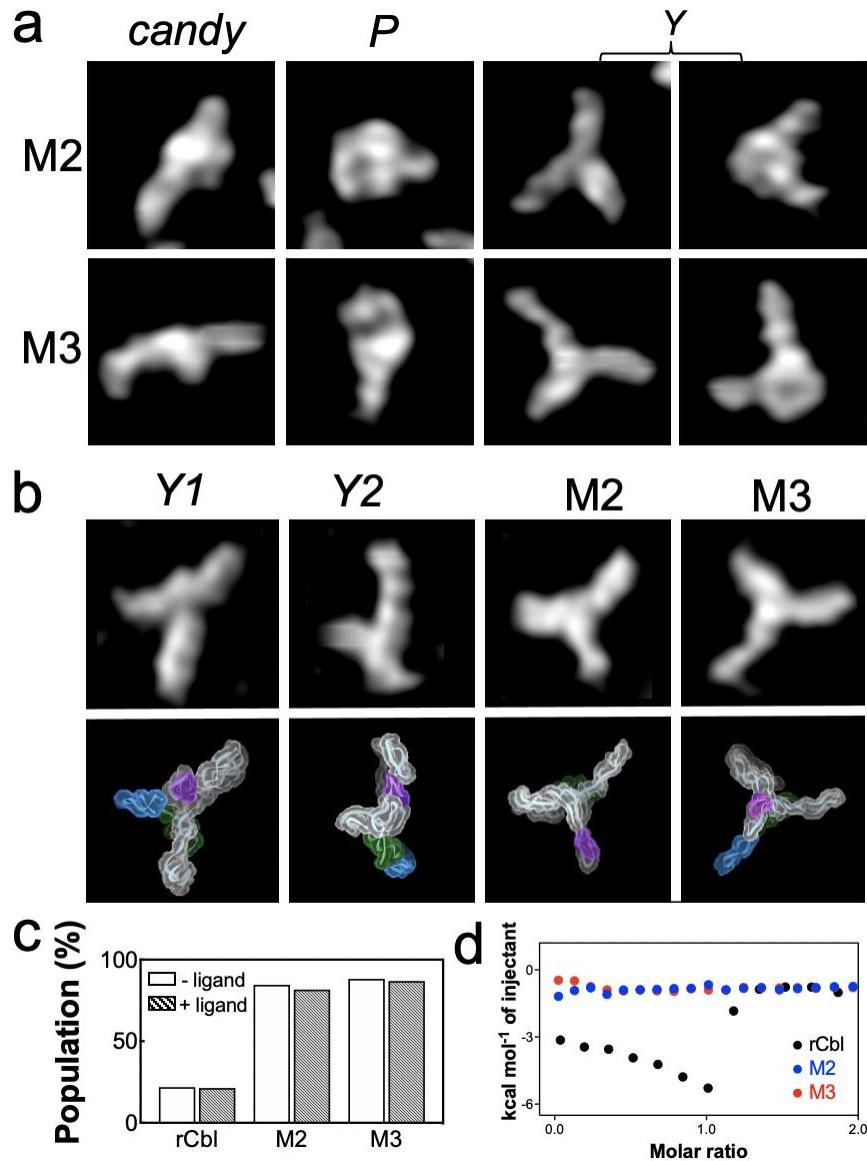


Figure 4 | Effects of KL interaction on rCbl folding. (a) Major conformations for rCbl M2 and M3 in absence of ligand. Both mutants exhibit dominant (>80%) *Y*-shaped particles, and minor (<10% each) *P*- and *candy*-shaped particles reminiscent of the non-mutant conformations. (b) AFM images (top) and recapitulated structures (bottom) for *Y1* and *Y2* of the wild-type rCbl, and M2 and M3 mutants. P1, P5 and P13 helices are colored in green, purple and blue, respectively. (c) Population tallies of *Y* conformers in the rCbl, M2 and M3 samples in absence or presence of ligand. In all cases, the *Y* population is unaffected by ligand binding. (d) Raw isotherms for rCbl (black) and M3 (red).

Methods

RNA sample preparation

All RNA samples were generated by *in vitro* transcription using synthesized linear DNA templates (IDT, Coralville, Iowa). **Supplementary Fig. 2** illustrates the extensive investigation of the factors that affect the conformations of the RNA samples and found that the presence of heterogeneous conformers in all cases, albeit with different populations depending on preparation protocols: salt concentration, temperature, storage conditions, the presence/absence of ligand and RNA concentration. It is noteworthy that the sample incubated at room temperature under high salt buffer conditions with high RNA concentration folds into almost homogeneous dimers (**Supplementary Fig. 2a, 2d**). In conclusion, we selected the protocol, as described below, which produces the sample with the highest population of binding-competent conformers based on the ITC data and AFM images as described in the following. The transcript was purified by 6% polyacrylamide gel electrophoresis under native conditions. The RNA was eluted from the gel using RNA elution buffer containing 50 mM sodium acetate, pH 5.3, 2 mM EDTA, followed by buffer exchange at least three times to high salt buffer (HSB) (10 mM HEPES, pH 7.5, 100 mM KCl, 0.1 mM EDTA, 10 mM MgCl₂).

AFM experiments

All AFM experiments were performed under solution conditions using a Cypher VRS AFM (Asylum Research, Oxford Instrument) at 4 °C with amplitude-modulated AC mode (commonly known as tapping mode). To immobilize RNA, mica supports were treated with 1-(3-aminopropyl) silatrane (APS). 50 mM APS stock was diluted 300-fold in water just before use and coated on freshly cleaved muscovite mica (Grade V1) (Ted Pella Redding, CA). After 30 min, the mica surface was rinsed with pico-pure water and dried gently with filtered nitrogen gas. For the sample

in absence of ligand, 10 μ l 20 nM rCbl was deposited on APS-mica for ~20 min and washed with 200 μ l low salt buffer (LSB). For the sample in presence of ligand, 50 nM rCbl RNA was mixed with 1 mM AdoCbl (coenzyme B12, Sigma-Aldrich, USA) ligand first, then 10 μ l mixed sample was put on APS-mica for ~15 min and washed with 200 μ l ligand buffer (50 mM MES, pH 6.0, 10 mM KCl, 1 mM MgCl₂, 1 mM AdoCbl). To get high-resolution images, FASTSCAN-D-SS probes (Bruker) were used, which had a nominal tip radius of 1 nm, a resonance frequency of 80~140 kHz and a normal spring constant of 0.25 N/m. A pulsed blue laser (BlueDrive) was used for photothermal excitation, positioned at the base of the cantilever, while a superluminescent diode (SLD) was positioned near the head of the cantilever to detect cantilever deflection. Images were collected with a scan size of 500 x 500 nm², 1024 x 1024 pixels² and a scan rate of 1.0 Hz, with a typical initial set point of 450 mV and a free amplitude of 500 mV. The setpoint was changed after the tip approach and adjusted during the imaging based on the image quality. In total, 11 and 9 high-quality images were collected for rCbl in the absence or presence of ligand, respectively.

AFM image processing and analysis

For images used for 3D topological structure calculation, the raw images were first processed through the following built-in settings in SPIP (Scanning Probe Image Processor) software: plane correction by applying 3rd-order polynomial leveling to the particle-free region, filtering by despiking to remove some artifact lines, and FFT (Fast Fourier Transform) analysis to remove high-frequency noise. The final image resolution was increased to 4096 x 4096 pixels² by doubling the number of pixels twice. Single-particle images were cropped from the processed images and converted to pseudoAFM (*.txt) with 5 Å resolution in MountainsSPIP for structure calculation.

The grayscale images were processed using Gwyddion (version 2.56). Horizontal streaks correction was first applied to remove streaks, followed by 3rd-order polynomial leveling to flatten the background. 2D FFT filtering was further applied to remove high-frequency noise. The final image digital resolution was increased to 4096 x 4096 pixels².

Tally analysis was done using the particle analysis function implemented in SPIP by manually selecting the particles for each class based on the topological surface. Particles with multimers or with linear nucleic acids were not counted.

3D topological structure calculations

A detailed description of recapitulation of 3D topological structures calculations from topographic AFM images is presented elsewhere. The following is a brief introduction to the method. We applied coarse-grained dynamic fitting to individual AFM particle topographic surfaces. The fitness of the model structures to the AFM molecular surfaces are examined by a cross-correlation score together with energetic considerations of the RNA hierarchical folding principle by applying a series of filters to select the top structures, followed by conversion to an all-atom atomic model using RS3D¹⁹ and all-atom refinement with Xplor-NIH²⁰.

Isothermal titration calorimetry (ITC) and data analysis

ITC experiments were performed using a MicroCal iTC200 (Malvern, United Kingdom) at 25 °C, following a standard procedure of RNA-ligand titration as described previously²¹. RNA sample was buffer exchanged extensively to ITC buffer containing 50 mM MES, pH 6.0, 10 mM KCl, 1 mM MgCl₂ and coenzyme B12 (AdoCbl) ligand was dissolved in ITC buffer. 70 μM RNA was titrated with 700 μM AdoCbl. The blank experiment was performed by replacing RNA sample with ITC buffer and was subtracted from the experimental data during data analysis. The baseline

correction and integration of the thermogram were processed using NITPIC²². The raw thermogram was deconvoluted into two principal components by singular value decomposition analysis (SVD) using a MATLAB script. The thermodynamic parameters for each component were determined by nonlinear regression fitting using either a one-site or two-site binding model. The raw thermogram of AdoCbl binding to rCbl appeared unconventional, and differ from a typical exothermic binding heat compensation followed by an endothermic response (**Supplementary Fig. 8a**), suggesting the AdoCbl binding triggers a substantial conformational change in the riboswitch aptamer *via* an induced-fit binding scenario, which is further evidenced by our AFM tally results (main text) in the presence and absence of AdoCbl.

To determine the number of states associated with the AdoCbl binding reaction and rationally reconstruct the ITC thermogram with all significant components (**Supplementary Fig. 8b-d**), a series of titration peaks of the ITC thermogram (**Supplementary Fig. 8a**) were subjected to SVD analysis using MATLAB (MATLAB and Statistics Toolbox Release 2021a, The MathWorks, Inc., Natick, Massachusetts, United States). The heat compensation profiles related to each titration point were used to produce an $m \times n$ matrix, M, and used as input for SVD analysis, where m corresponds to the number of the recorded heat compensation profiles and n corresponds to the number of titration points at a different ligand to RNA ratios. To unbiasedly determine the number of significant components (**Supplementary Fig. 8e**), the normalized autocorrelation coefficients of individual singular values were calculated with a minimum threshold of 0.75 as a selection filter (dashed line in **Supplementary Fig. 8f**). Those SVD components with autocorrelation coefficients less than 0.75 were systematic noise and were not used to reconstruct the ITC thermogram and to derive the related thermodynamic parameters. The isotherms integrated from selected thermograms were subjected to nonlinear regression fitting by the 1:1 one-site binding model or

two-site independent binding model, as described previously²³, and fitting results referring to each SVD component were tabulated (**Supplementary Table 2**). The dissociation constants, binding stoichiometry, and thermodynamic parameters of AdoCbl binding to rCbl were derived by fitting the deconvoluted isotherms to the numerical solution of the explicit ordinary differential equation (ODE) of Wiseman isotherm formulation^{23,24} using MATLAB and GraphPad Prism (GraphPad Software, La Jolla, California, USA). The choice of model, *i.e.*, one-site, or two-site binding model, was determined using the F-test.

To confirm that the mutant Y conformers were inactive, the same ITC experiments were conducted for both rCbl and M3 on the same day under the same solution conditions and experimental settings, except for RNA and ligand concentrations, which were 50 μ M RNA and 500 μ M, respectively.

Small angle X-ray scattering (SAXS) experiments and data analysis

The SAXS experiments were carried out at the 12ID-B beamline of the Advanced Photon Source (APS), Argonne National Laboratory. Photon energy was 13.3 keV ($\lambda=0.932 \text{ \AA}$). A sample-to-detector distance of 1.9 m was used to obtain q range of $0.005 < q < 0.88 \text{ \AA}^{-1}$. Here, q is the magnitude of the scattering vector, $q = (4\pi/\lambda)\sin\theta$, where 2θ is the scattering angle and λ is the wavelength of the radiation. Ideally, the concentration of rCbl for SAXS should be the same as for the AFM experiments. However, the concentration of 20 nM for AFM is too low to produce reasonable SAXS signals. To select a concentration that is as low as possible and yet produces reasonable scattering curves after background subtraction, a wide range of RNA concentrations from 25 nM to 2 μ M were analyzed. For RNA samples complexed with ligand, the concentrations of rCbl were in the same range of 25 nM to 2 μ M, with the molar ratio of rCbl to ligand of 1:10.

Buffers containing ligand at concentrations in the range of 225 nM to 18 μ M were measured, accordingly, and used for background subtraction for ligand-containing samples. To minimize radiation damage and obtain a good signal-to-noise ratio, 225 images were taken for each sample and buffer using a flow cell with an exposure time of 1 second. The two-dimensional scattering patterns were collected using a Pilatus 2M detector and converted to one-dimensional SAXS curves through radial averaging after solid angle correction and then normalizing with the intensity of the transmitted X-ray beam. Data were then averaged after the elimination of outliers using the software package developed at beamline 12-ID-B.

The radius of gyration R_g and intensity at angel zero, $I(0)$, were generated from the Guinier plot in the range of $qR_g < 1.3$. For comparison, R_g and $I(0)$ were also calculated in real and reciprocal spaces using the program GNOM in the q range up to 0.30 \AA^{-1} ²⁵. The pair-distance distribution function $P(r)$ and maximum dimension (D_{\max}) were also calculated using GNOM. The molecular weights were estimated based on the method of correlation volume, V_c , using the formula for RNA²⁶. The obtained structural parameters and molecular weights are listed in **Supplementary Table 3**.

The SAXS data for rCbl at the concentration of 1 μ M were selected for the ensemble fitting. The synthesized scattering intensity from an ensemble of conformations can be described as:

$$I_{syn}(q_i) = \sum_{k=1}^{N_{ens}} I_{calc}^k(q_i) \cdot v_k$$

where v_k and $I_{calc}^k(q)$ are the volume fraction and calculated scattering intensity from the k^{th} component/conformation, respectively. N_{ens} is the number of conformations in the ensemble. All conformations used for fitting were obtained by AFM. The scattering profiles for each conformation in the ensemble were calculated in the q range $0 < q < 0.50 \text{ \AA}^{-1}$ using Crysolv 3.0²⁷.

The volume tallies from AFM (**Supplementary Fig. 3**) were used to generate initial values of volume fractions ν_k for each conformation.

The goodness-of-fit of the synthesized scattering intensity from an ensemble to the SAXS experimental data was evaluated by comparing the synthesized profile, $I_{syn}(q)$, with the experimental one $I_{saxs}(q)$.

$$\chi^2 = \frac{1}{N_q - 1} \sum_{i=1}^{N_q} \left[\frac{I_{saxs}(q_i) - a \cdot I_{syn}(q_i) - b}{\sigma(q_i)} \right]^2$$

where, $a = \sum_{i=1}^{N_q} I_{saxs}(q_i) \cdot I_{syn}(q_i) / \sum_{i=1}^{N_q} I_{saxs}(q_i) \cdot I_{saxs}(q_i)$, b is background offset, N_q is a number of experimental points, $\sigma(q)$ is the experimental error.

Data availability

All structural model coordinates and computation files can be downloaded at <http://...>

References

1. Kendrew, J.C. et al. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-6 (1958).
2. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
3. Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590-596 (2021).
4. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876 (2021).
5. Ganser, L.R., Kelly, M.L., Herschlag, D. & Al-Hashimi, H.M. The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol* **20**, 474-489 (2019).
6. Larsen, K.P., Choi, J., Prabhakar, A., Puglisi, E.V. & Puglisi, J.D. Relating Structure and Dynamics in RNA Biology. *Cold Spring Harb Perspect Biol* **11**(2019).
7. Al-Hashimi, H.M. & Walter, N.G. RNA dynamics: it is about time. *Curr Opin Struct Biol* **18**, 321-9 (2008).
8. Tomezsko, P.J. et al. Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **582**, 438-442 (2020).
9. Townshend, R.J.L. et al. Geometric deep learning of RNA structure. *Science* **373**, 1047-1051 (2021).
10. Schon, P. Atomic force microscopy of RNA: State of the art and recent advancements. *Semin Cell Dev Biol* **73**, 209-219 (2018).
11. Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biology* **8**, R239 (2007).
12. Gardner, P.P. et al. Rfam: updates to the RNA families database. *Nucleic Acids Res* **37**, D136-40 (2009).

13. Rodionov, D.A., Vitreschak, A.G., Mironov, A.A. & Gelfand, M.S. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J Biol Chem* **278**, 41148-59 (2003).
14. Vitreschak, A.G., Rodionov, D.A., Mironov, A.A. & Gelfand, M.S. Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *RNA* **9**, 1084-97 (2003).
15. Holmstrom, E.D., Polaski, J.T., Batey, R.T. & Nesbitt, D.J. Single-molecule conformational dynamics of a biologically functional hydroxocobalamin riboswitch. *J Am Chem Soc* **136**, 16832-43 (2014).
16. Choudhary, P.K. & Sigel, R.K. Mg²⁺-induced conformational changes in the btuB riboswitch from *E. coli*. *RNA* **20**, 36-45 (2014).
17. Chan, C.W. & Mondragon, A. Crystal structure of an atypical cobalamin riboswitch reveals RNA structural adaptability as basis for promiscuous ligand binding. *Nucleic Acids Res* **48**, 7569-7583 (2020).
18. Johnson, J.E., Jr., Reyes, F.E., Polaski, J.T. & Batey, R.T. B12 cofactors directly stabilize an mRNA regulatory switch. *Nature* **492**, 133-7 (2012).
19. Bhandari, Y.R. et al. Topological structure determination of RNA using Small-Angle X-Ray Scattering. *J Mol Biol* **429**, 3635-3649 (2017).
20. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. & Clore, G.M. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65-73 (2003).
21. Jones, C.P., Piszczek, G. & Ferre-D'Amare, A.R. Isothermal Titration Calorimetry Measurements of Riboswitch-Ligand Interactions. *Microcalorimetry of Biological Molecules* **1964**, 75-87 (2019).
22. Keller, S. et al. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem* **84**, 5066-73 (2012).
23. Poon, G.M. Explicit formulation of titration models for isothermal titration calorimetry. *Anal Biochem* **400**, 229-36 (2010).

24. Wiseman, T., Williston, S., Brandts, J.F. & Lin, L.N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179**, 131-7 (1989).
25. Svergun, D.I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *Journal of Applied Crystallography* **25**, 495-503 (1992).
26. Rambo, R.P. & Tainer, J.A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477-81 (2013).
27. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212-1225 (2017).

Acknowledgments This work was funded by the Intramural Research Program of the National Cancer Institute, National Institutes of Health (Y-X.W.) and in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN261201500003I (R.N.). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We thank Will F. Heinz for technical assistance in AFM, Benjamin Miller, Steve Fellini and Susan Chacko of the NIH HPC computing facility for allocating computing and storage resources for the project, and Dr. Jeffrey N. Strathern for his vision about RNA biology.

Author contributions J.D. and P.Y. prepared samples; J.D. recorded all AFM images; J.D. and Y-X.W. performed all calculations of the structures; Y.R.B and C.D.S. wrote all computation codes; L.F. recorded and analyzed SAXS data; L.F. and Y-X.W. interpreted experimental SAXS data; J.D., Y-T.L. and S.G.T. recorded ITC data; Y-T.L. analyzed/interpreted the ITC data; B.M and R.N. participated in the initial phase of the project; J.R.S. for very careful examination of the electron density map of the rCbl crystal structure; A.R. and J.Z. for many insightful discussion; Y-X.W. conceptualized, designed the project, drafted the manuscript; all authors contributed to the revision.

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.-X.W. (wangyunx@mail.nih.gov)

SUPPLEMENTARY INFORMATION

Visualizing RNA conformational and architectural heterogeneity in solution

Jienyu Ding¹, Yun-Tzai Lee¹, Yuba Bhandari¹, Lixin Fan³, Charles D. Schwieters², Ping Yu¹, Sergey G. Tarosov⁴, Jason R. Stagno¹, Buyong Ma⁵, Ruth Nussinov⁵, Alan Rein⁶, Jinwei Zhang⁷, Yun-Xing Wang¹

¹ Protein-Nucleic Acid Interaction Section, Center for Structural Biology, National Cancer Institute, Frederick, Maryland 21702, USA; ² Computational Biomolecular Magnetic Resonance Core, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA; ³ Basic Science Program, Frederick National Laboratory for Cancer Research, Small Angle X-ray Scattering Core Facility of National Cancer Institute , National Institutes of Health, Frederick, Maryland 21702, USA ⁴ Biophysical Resource, Center for Structural Biology National Cancer Institute, Frederick, Maryland 21702, USA; ⁵ Computational Structural Biology Section, Frederick National Laboratory for Cancer Research in the Laboratory of Cancer Immunometabolism, National Cancer Institute, Frederick, MD 21702, USA; ⁶ Retrovirus Assembly Section, HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702, USA; ⁷ Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA

Corresponding author: Yun-Xing Wang, wangyunx@mail.nih.gov

Supplementary discussion

The effects of salt concentration on RNA conformational heterogeneity and activity

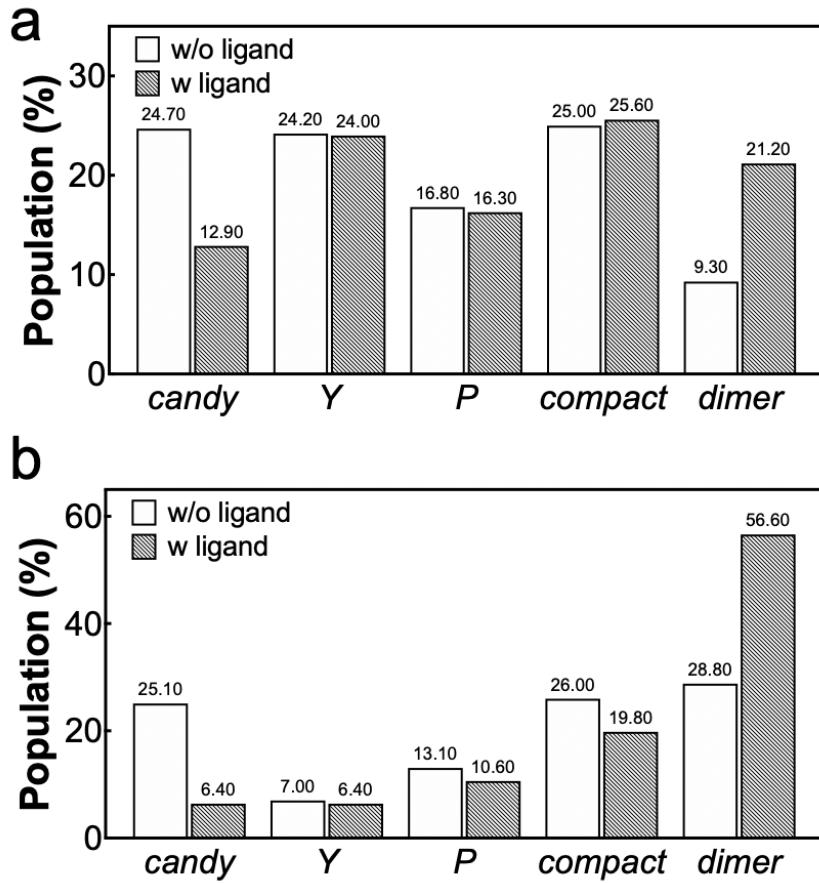
The high-resolution AFM images were obtained under low salt buffer (LSB) conditions (10 mM KCl, 1 mM MgCl₂). We found, however, that when rCbl was prepared in LSB after gel elution, and stored under these conditions until AFM imaging, the majority of RNA molecules exhibited *Y* conformations and showed very little differences in the absence or presence of ligand (**Extended Data Fig. 1a-b**). We suspect that, although LSB conditions are optimal for imaging, they are not suitable for maintaining rCbl folding stability over several hours or days. For this reason, rCbl samples were kept in high salt buffer (HSB) conditions (100 mM KCl, 10 mM MgCl₂) until just before AFM imaging, at which point they were diluted in LSB and used immediately. Despite rCbl prepared in LSB being predominantly *Y* conformers by AFM, the RNA was still active by ITC analysis, possibly due to the effects of much higher RNA and ligand concentrations (discussed below). However, the ITC data for LSB- and HSB-prepared samples did reveal differences in thermodynamic behavior (compare **Fig. 2e, 4d**, and **Extended Data Tables 2, 5**).

The effects of RNA and ligand concentrations on conformational heterogeneity

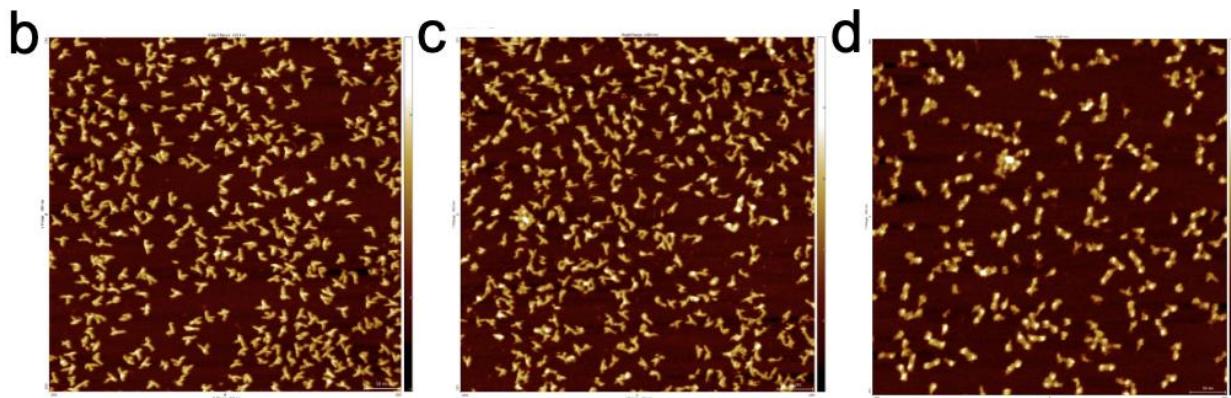
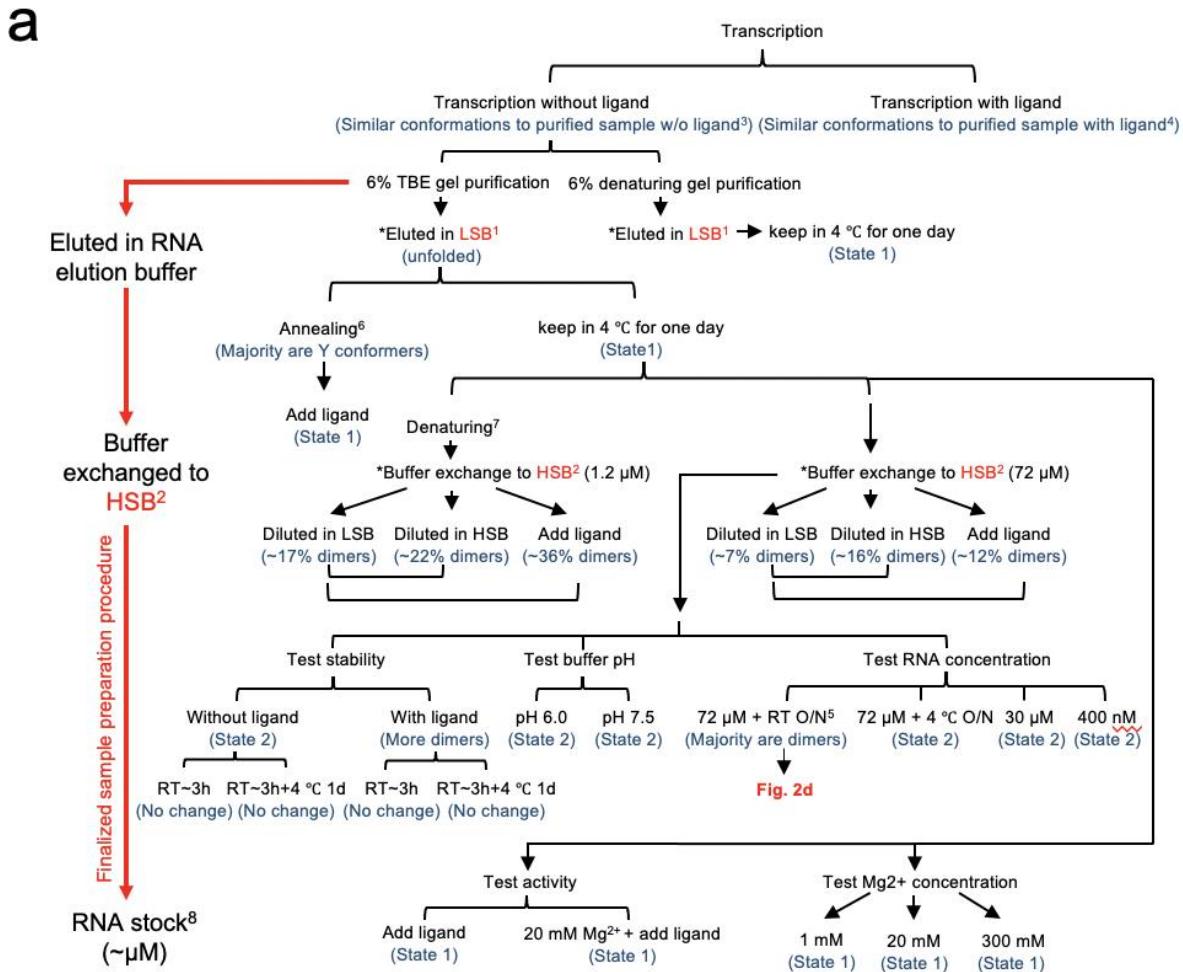
Given the number of various monomer and dimer conformations observed by AFM, the convolution of heat exchange upon ligand binding is likely of greater complexity than can be explained from ITC data, particularly with regard to the effects of RNA and ligand concentrations used in each method, and conformational heterogeneity under equilibrium conditions prior to ligand binding. For instance, at much lower (nM) RNA concentration, as indicated by AFM imaging, conformational changes and dimer formation comprise the bulk of the observed differences induced by ligand binding. Whereas, at the much higher (μ M) ITC concentration, the majority of the RNA adopts a bound-like state prior to ligand binding. Therefore, the extent of induced-fit versus conformational selection mode of binding, as it relates to the binding-competent population, ligand-driven conformational changes, and dimer formation, may vary with

RNA and ligand concentrations. As such, ITC-, SAXS-, and AFM-derived conformer populations, which are recorded at different RNA and ligand concentrations, are not identical, even though they exhibit consistent trends.

Extended Data Figures



Extended Data Figure 1 | Population tallies for rCbl ITC and SAXS samples. AFM imaging was performed on the exact samples used for ITC (a) and SAXS (b) measurements for a more accurate comparison. Aliquots of these samples were taken before and after ligand addition, diluted to 20 nM, imaged by AFM, and tallied accordingly. The stock concentrations for ITC and SAXS samples were 70 μ M and 2 μ M, respectively.



Extended Data Figure 2 | Extensive study on factors affecting the conformations of rCbl RNA. (a) Illustration of sample preparation protocols with corresponding notes.

¹LSB: 50 mM MES, pH 6.0, 10 mM KCl, 1 mM MgCl₂.

²HSB: 10 mM HEPES, pH 7.5, 100 mM KCl, 0.1 mM EDTA, 10 mM MgCl₂.

³Conformations in purified sample without ligand: mixture of *Y*, *candy*, and *P* monomers (~98%), and various dimers (~2%) (see **Fig. 1a** in main text).

⁴Conformations in purified sample with ligand: mixture of *Y*, *candy*, and *P* monomers (~37%), and various dimers (~63%) (see **Fig. 1b** in main text).

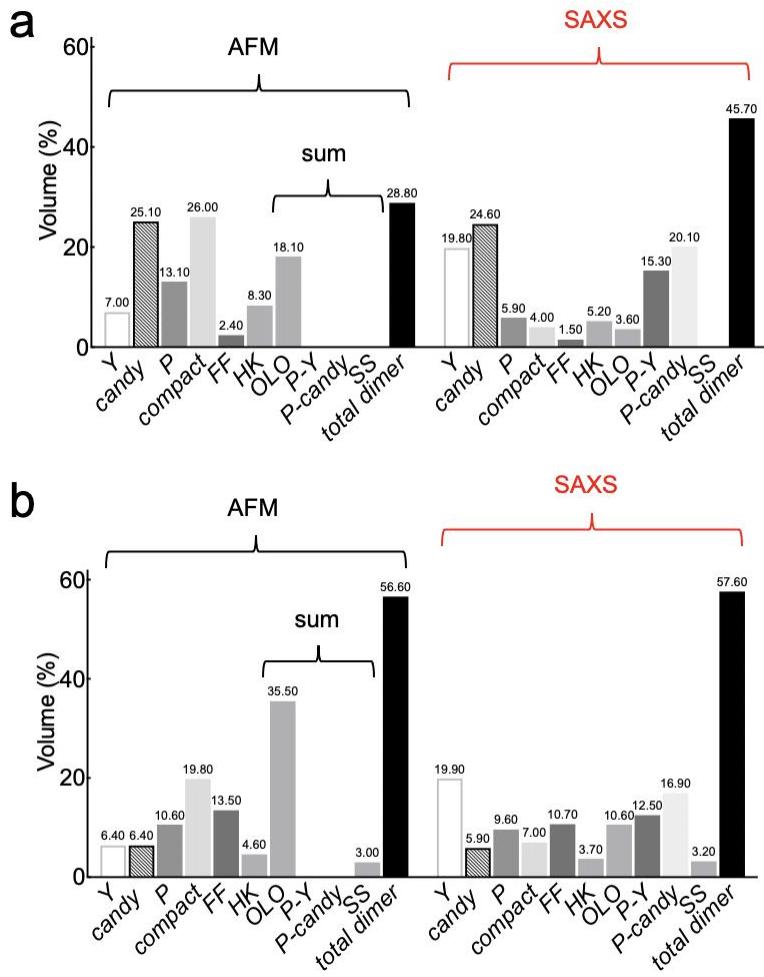
⁵Incubating 72 μ M RNA sample under high salt buffer at room temperature overnight results in much more homogeneous conformations (see **Extended Data Fig. 2d**).

⁶Annealing: sample heated at 85 °C for 2 min, followed by snap-cooling on ice.

⁷Denaturing: sample mixed with 8 M urea solution (DEPC H₂O + 1 mM EDTA), and heated at 65 °C for 5 min.

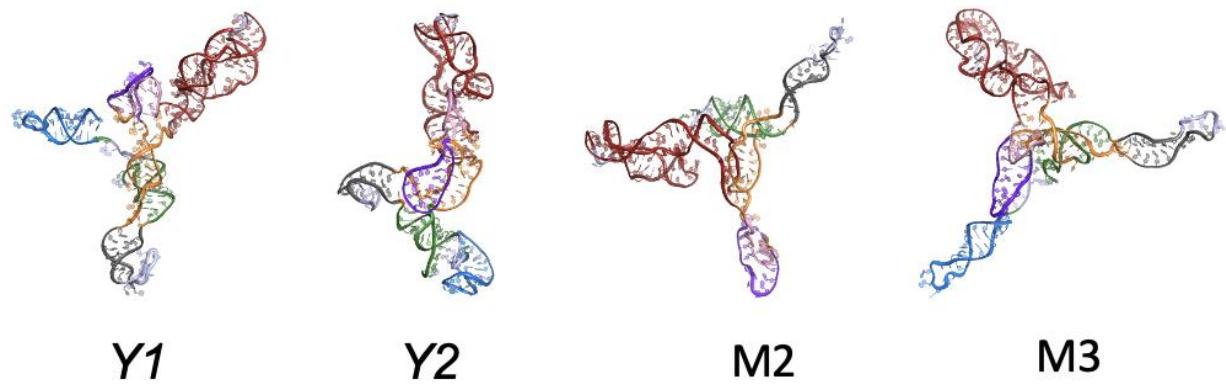
⁸Finalized sample preparation procedure for rCbl. All samples presented in main text were prepared in this way. Elution and buffer exchange steps marked with * result in concentration of RNA. (**b-d**)

The AFM images for State 1 (**b**), State 2 (**c**), and homogeneous dimers (**d**).

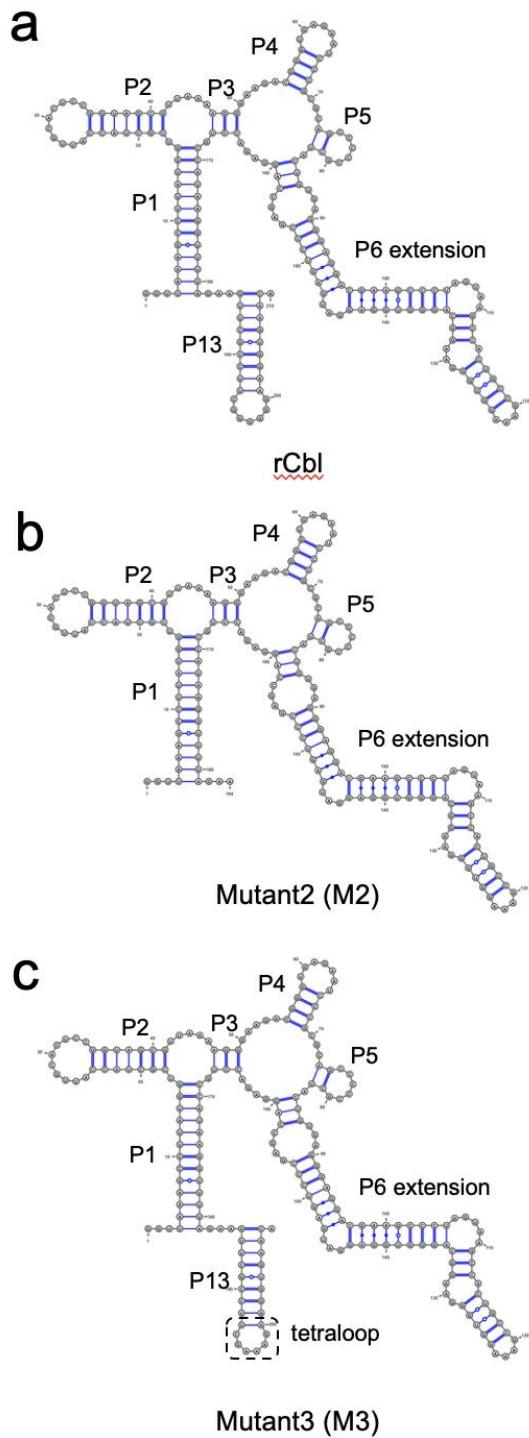


Extended Data Figure 3 | The volume percentages of the AFM-tallied and SAXS-derived conformers.

The volume percentages of monomeric and dimeric conformers in the absence (**a**) or presence (**b**) of ligand tallied from AFM images (left) or derived from ensemble fitting of SAXS data using AFM-calculated structures (right). The AFM-tallied volume percentages were calculated based on the volumes of representative conformers (**Extended Data Table 4**) and the population tallies of each class (**Fig. 1e**). The volume percentages of the SAXS-derived ensembles represent the linear combinations of conformers that yielded the lowest χ^2 values (no ligand:1.68, with ligand:1.00) (**Fig. 2c, d**). The samples used for AFM imaging were diluted aliquots of the exact samples used for SAXS measurements.

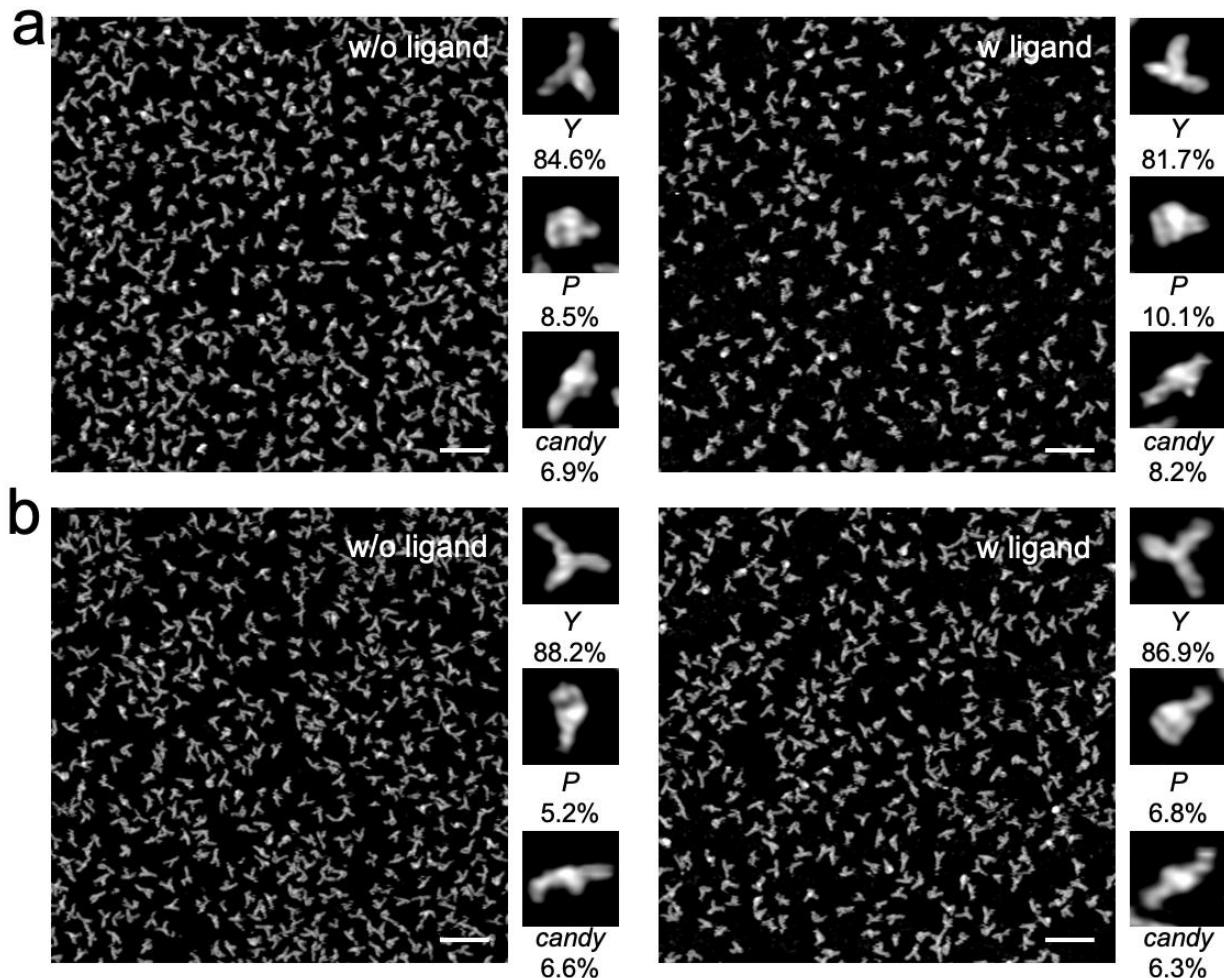


Extended Data Figure 4 | Recapitulated structures in the cartoon model of the Y conformers. P1, P2, P3, P4, P5, P6 extension, P13 are colored in green, gray, orange, purple, pink, red, and blue, respectively.

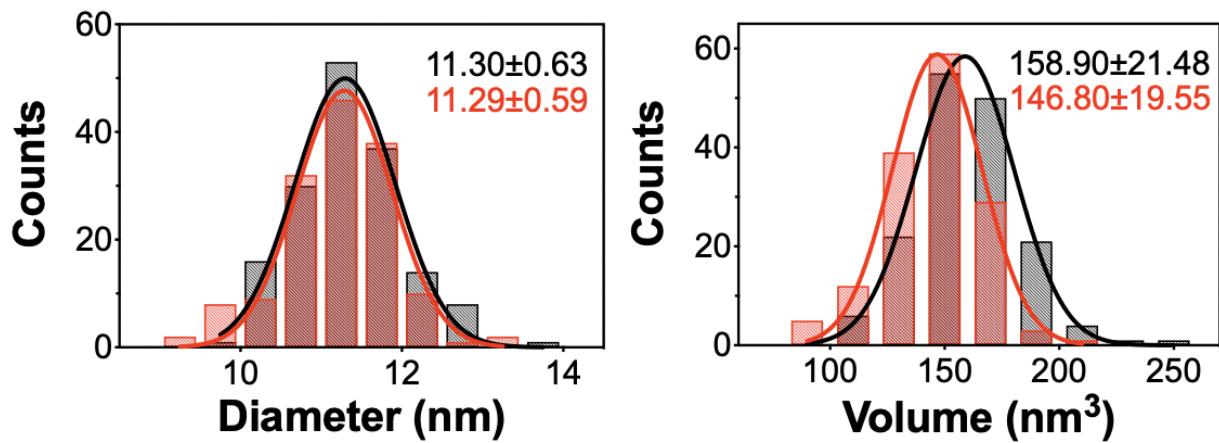


Extended Data Figure 5 | The secondary structures of the wild-type rCbl and two mutants. The wild-type (a), M2 (b) and M3 (c) mutants. In M2, the P13 helix is deleted entirely. In M3, L13

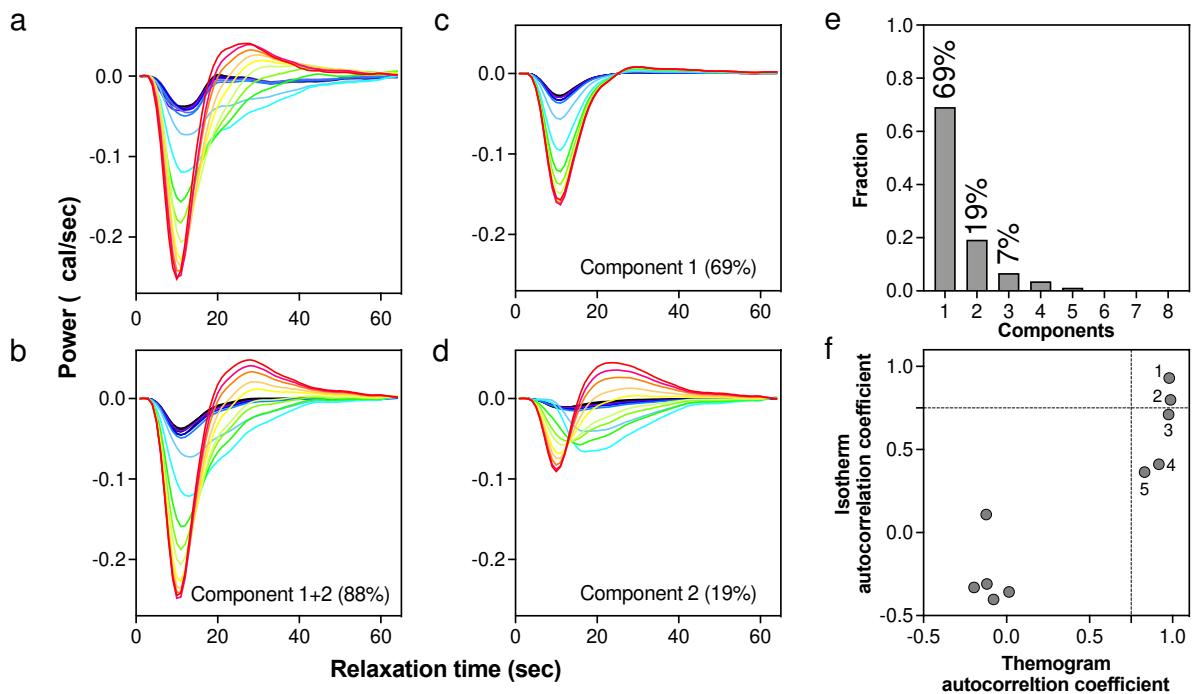
(GGGGG) is replaced with a tetraloop (GAAA). Both mutants are designed to disrupt the L5-L13 KL interaction.



Extended Data Figure 6 | AFM imaging of rCbl mutants that abrogate the KL interactions. AFM images of M2 (**b**) and M3 (**b**) in the absence (left) or presence (right) of 1 mM AdoCbl. The scale bars showing in the right bottom corner of each image represent 50 nm. The same three types of conformations (*Y*, *P*, *candy*) with nearly identical populations were observed in all images. Insets show representative particles for each conformer type and their respective populations tallies.



Extended Data Figure 7 | Particle diameter and volume histograms. Particle diameters (left) and particle volumes (right) for Y conformers of M2 (black, 160 particles) and M3 (red, 148 particles) determined (SPIP, Metrology). The Gaussian fits with mean values and standard deviations are shown.



Extended Data Figure 8 | SVD analysis of the ITC thermogram. **(a)** The raw ITC thermogram was deconvoluted and reconstructed using a minimum of two principal components determined by SVD. The heat compensation profiles were color-ramped from red to deep blue corresponding to each titration point. **(b)** The SVD-reconstructed thermogram corresponding to the two major components represents 88% of the total singular values. The SVD-deconvoluted thermograms for Components 1 and 2 are shown in panels **(c)** and **(d)**, respectively. **(e)** The percentage of singular values and **(f)** their related autocorrelation coefficients determined by SVD analysis.

Extended Data Tables

Extended Data Table 1. Statistics table for rCbl topological structure calculations for each monomer and dimer class.

	Rg	E(total)	E(covalent)	E(tertiary)	E(stack)	E(HB)	E(repul)	E(elect)	E(AFM)	aftmCC
compact	30.79	-3896.85	568.5	-464.34	-2589.01	-1984.18	1.93	235.29	333.97	0.9942886
candy	38.56	-231.35	489.78	-623.03	-221.03	-181.97	2.98	216.49	84.45	0.9924422
candy2	39.24	-4085.23	494.14	-368.7	-2563.82	-1935.54	2.51	183.32	101.87	0.9875678
P	32.02	-3178.03	562.81	-595.13	-1994.8	-1516.29	0.97	217.67	145.76	0.9869547
FF	58.75	-20481.2	1201.69	-6036.4	-4388.38	-12587.58	12.3	512.28	803.9	0.9937717
OLO	59.71	-9200.1	1073.7	-3102.86	-3879.11	-4385.09	12.66	479.47	600.14	0.9884903
SS	44.47	-26044.27	1161	-8601.29	-5219.31	-15916.64	58.5	585.19	1898.73	0.9494404
PY	56.94	-16979.12	1061.31	-5234.3	-3856.45	-9786.39	10.28	424.38	401.06	0.9907171
Y1	45.02	-3614.53	611.7	-240.09	-2412.63	-1908.31	1.6	179.64	152.56	0.9918077
Y2	44.59	-5091.41	595.01	-216.66	-3100.2	-2635.08	2.52	169.82	92.2	0.9973666
M21	44.94	-2750.53	483.53	-191.51	-1948.34	-1975.28	1.66	131	747.47	0.930575
M31	49.12	-4377.74	545.91	-265.68	-1931.64	-2998.29	3.08	157.79	110.12	0.9729926

Extended Data Table 2. Isothermal titration calorimetry (ITC)-derived thermodynamic parameters for the interaction of rCbl with AdoCbl after singular value decomposition (SVD) under LSB (**Fig. 2e**).

		K_d (nM)	N_{app}	ΔG (kcal/mol)	ΔH (kcal/mol)	-ΔST (kcal/mol)
component1	69%	626.2±40.5	0.76*	-8.46	-1.69±0.04	-6.77
		600.6±69.1	0.78	-8.56	-1.68±0.03	-6.80
component2	19%	1347±266 [†]	0.65*	-7.85	4.03±0.24	-12.03
		1738±510 [†]	0.71	-7.85	4.46±1.15	-12.31
		85.0±54.4 [†]	0.88*	-9.64	-1.69±0.13	-7.95
		136.4±120 [†]	0.89	-9.36	-2.90±1.15	-6.64

* Stoichiometric N_{app} (component1) and mean N_{app} (component2) value of 0.76 used for fitting based on the AFM fraction of conformers in the absence of ligand that were not Y-shaped (76%) (**Extended Data Fig. 1a**).

† Standard error greater than 10% of the mean.

Extended Data Table 3. Molecular weights estimated from experimental SAXS data.

	R _g Guinier (Å)	R _g real space (Å)	D _{max} (Å)	MW (V _c)*	MW (fit) [†]	Monomer MW (kDa)
rCbl (no ligand)	58.89±2.41	61.1±1.13	207	86.7	99.8	68.5
rCbl (1mM ligand)	57.10±2.15	58.99±1.10	197	97.7	108.0	70.0

* MW (V_c): Mean particle weight derived from experimental SAXS data using method based on correlation volume for RNA at Q_{max}=0.3 Å⁻¹.

† The MWs (fit) are calculated using the dimer and monomer volume percentages from the SAXS fitting results, plus the theoretical MWs for monomer and dimer, as the equation shown below:

MW(fit) = dimer volume % x theoretical dimer MW + monomer volume % x theoretical monomer MW
(Extended Data Fig. 3).

Extended Data Table 4. Volume and radius of gyration (R_g) values for various conformers.

	Y	candy	P	compact	FF	OLO	SS	P-Y	P-candy
Volume (nm ³)	49.38	49.39	50.88	49.32	101.64	101.56	101.11	101.81	101.06
R _g (Å)	44.84	38.74	38.60	34.21	57.42	62.54	44.27	54.55	61.99

Note:

The volume values were assessed using 3V Web Server (<http://3vee.molmovdb.org/>)¹ by setting the probe radius to 0. R_g values were calculated using CRYSTOL²

Extended Data Table 5. Isothermal titration calorimetry (ITC)-derived thermodynamic parameters for the interaction of rCbl with AdoCbl after singular value decomposition (SVD) under HSB.

	K _d (nM)	N _{app}	ΔG (kcal/mol)	ΔH (kcal/mol)	-ΔST (kcal/mol)
component1	71%	289.2±76.4	0.855*	-8.91	-4.22±0.13
component2	14%	681.5±236.0 [†]	0.700*	-8.41	4.86±0.23
		32.8±16.3 [†]	1.010*	-10.20	-4.75±0.25
					-5.45

* Stoichiometric N_{app} (component1) and mean N_{app} (component2) value of 0.855 used for fitting based on the AFM fraction of conformers in the absence of ligand that were not Y-shaped (85.5%) (data not shown here).

† Standard error greater than 10% of the mean.

Reference

- 1 Voss, N. R. & Gerstein, M. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res* **38**, W555-562, doi:10.1093/nar/gkq395 (2010).
- 2 Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212-1225, doi:10.1107/S1600576717007786 (2017).