Biophysical and dynamic characterization of a fine-tuned binding of the human Respiratory Syncytial Virus M2-1 core domain to long RNAs

Icaro P. Caruso[1,2,3], Giovana C. Guimarães[1], Vitor B. Machado[1], Marcelo A. Fossey[1,2], Dieter Willbold[4], Fabio C. L. Almeida[3,*], Fátima P. Souza[1,2,*]

[1]Multiuser Center for Biomolecular Innovation (CMIB), [2]Department of Physics, IBILCE/UNESP, 15054-000, São José do Rio Preto, SP, Brazil;

[3]National Center for Nuclear Magnetic Resonance of Macromolecules, Institute of Medical Biochemistry Leopoldo de Meis (IBqM) and National Center for Structural Biology and Bioimaging (CENABIO), UFRJ, 21941-590 Rio de Janeiro, RJ, Brazil

[4]Institute of Biological Information Processing, Structural Biochemistry and JuStruct (IBI-7), Forchungszentrum Jülich, 52425 Jülich; Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße, 40225 Düsseldorf, Germany.

[*]Corresponding author e-mail address: falmeida@bioqmed.ufrj.br

fatima.p.souza@unesp.br

Tel.: +55-21-31042326

Tel.: +55-17-32212240

ABSTRACT

The human Respiratory Syncytial Virus (hRSV) M2-1 protein functions as a processivity and antitermination factor of the viral polymerase complex. Here it is presented the first evidence that hRSV M2-1 core domain (cdM2-1) alone has an unfolding activity for long RNAs, as well as a biophysical and dynamic characterization of the cdM2-1/RNA complex. The main contact region of cdM2-1 with RNA was the α1–α2–α5–α6 helix bundle, which suffered local conformational changes and promoted the RNA unfolding activity. This activity may be triggered by base-pairing recognition. RNA molecules wrap around the whole cdM2-1, protruding their terminals over the domain. The α2–α3 and α3–α4 loops of cdM2-1 were marked by an increase in picosecond internal motions upon RNA binding even though they are not directly involved in the interaction. The results revealed that the cdM2-1/RNA complex originates from a fine-tuned binding, contributing to unraveling interaction aspects necessary to M2-1 activity.

IMPORTANCE

The main outcome is the molecular description of a fine-tuned binding of the cdM2-1/RNA complex and the evidence that the domain alone has an unfolding activity for long RNAs. This binding mode is essential in the understanding of the function in the full-length protein. Orthopneumovirus, as the human Respiratory Syncytial Virus (hRSV), stands out for the unique role of M2-1 as a transcriptional antitermination factor able to increase the RNA polymerase processivity.

KEYWORDS

hRSV M2-1 core domain; RNA-binding protein; fine-tuned binding; RNA unfolding activity; NMR

INTRODUCTION

Human Respiratory Syncytial Virus (hRSV) is the major causative agent of lower tract respiratory diseases such as pneumonia and bronchiolitis in children worldwide. This virus is also responsible for respiratory morbidity in the elderly and people with compromised immunity and cardiorespiratory disease. The main hRSV-related risk group is children with congenital immunodeficiency, bronchopulmonary dysplasia, heart disease, hypertension, prematurity, and low birth weight (1–3).

Of the 11 proteins encoded by hRSV, two are expressed by the M2 gene, so called M2-1 and M2-2 (1). The proteins of the M2 gene are involved in the assembly of the active form of the ribonucleoprotein (RNP) complex, with the M2-1 acting in transcription as a processivity and antitermination factor, and M2-2 working as a molecular "switch" between transcription and RNA replication (4).

M2-1 prevents the RNA-dependent RNA polymerase (RdRp) complex from dissociating prematurely at intragenic sites and also allows it to read through gene and termination signals, thus increasing the transcription efficiency of genes near the 5' end (5, 6). Blondot and collaborators (2012) hypothesized three possible explanations for acting as a factor that prevents the premature termination of the transcription, thus important for mRNA transcription: i) The M2-1 protein could bind to the nascent mRNA transcript to favor the transcription elongation, either preventing it to re-hybridize to the template or from forming secondary structures, which would destabilize the transcription complex. ii) The polymerase processivity enhancing effect of M2-1 could be due to an increase in the affinity of the polymerase for the genomic RNA template in a sequence non-specific manner. iii) The M2-1 could recognize gene end sequences either on the nascent mRNA or on the RNA template, preventing the release of the polymerase complex from its template and favoring transcription re-initiation at the downstream gene start sequences (7).

The crystal structure of the hRSV M2-1 protein determined by Tanner and collaborators (2014) shows a tetrameric arrangement with the monomers presenting three structurally distinct regions: the zinc finger domain which is composed of N-terminal residues bind to a zinc atom and aids in the transcription process; the oligomerization helix that is responsible for tetramer formation; and the core domain that acts directly on the interaction with phosphoprotein P and RNA (8). The core domain of M2-1 (cdM2-1) is folded into six α-helices, which are structurally arranged in a α1–α2–α5–α6 helix bundle and α3–α4 hairpin (7, 8). Blondot et al. (2012) determined by NMR studies that α2, α5, and α6-helix of the cdM2-1 play a key role in the binding to nucleic acids, and also demonstrate the preference of the core domain for purine-rich RNAs, especially adenine (7).

The crystal structure of the human Metapneumovirus (hMPV) M2-1 protein, which is homologous to hRSV M2-1, presented a distinct conformation for one of the tetramer core domains, which was termed as open conformation (9). In this conformation, one of the tetramer core domains is located far from the rest of the protein due to a rotation in the flexible linker (residues with electronic density absent in the crystal structures) between the oligomerization and core domain. From molecular dynamics (MD) simulations and small-angle X-ray scattering (SAXS) experiments, it was characterized a dynamic open-closed conformation equilibrium for hMPV M2-1, and that the closed conformation is structurally similar to the hRSV M2-1 protein. It was also reported that M2-1 closed conformation is prevalent in the presence of RNA, which is stabilized by the simultaneous binding of the nucleic acid molecule to the zinc finger region and to the core domain of the protein (9).

Recently, Molina and coauthors (2018) observed that RNA of 20 nucleotides (20mer RNA) binds with positive cooperativity to two core domains in the hRSV M2-1 tetramer, allowing RNAs to bridge two adjacent RNA binding sites. No cooperativity was reported for the binding of the core domain to 20mer RNA and of M2-1 tetramer to 10mer RNA (10). They showed by circular dichroism (CD) that 20mer RNAs are unfolded upon the formation of its complex with

M2-1 tetramer but, in the presence of only core domain, significant structural changes are not detectable. CD approach indicated that the secondary structures content of M2-1 shows no major changes upon the formation of the complex (10). This finding corroborates with the first of the above-mentioned hypotheses proposed by Blondot and colleagues (2012) for the mechanisms by which M2-1 prevents the polymerase from terminating transcription (7). Molina and collaborators (2018) also reported by fluorescence experiments that the association process of nucleic acids with M2-1 tetramer is composed of two steps: the first being a fast phase related to conformational changes of the RNA molecules, and the second being a slow phase related to subtle rearrangements of the M2-1/RNA complex that occur through an induced-fit mechanism. Based on induced-fit mechanism, subtle conformational rearrangements in the complex take place in the protein moiety. Therefore, in the case of M2-1, as CD experiments indicate that no major secondary structure changes occur upon binding, probably loop or hinge motions may reposition the RNA binding domains (core domains), similarly as shown for the hMPV M2-1 protein, where transition from open to closed conformation is observed for cdM2-1 upon RNA association (9, 10).

Taken together, the aforementioned studies characterize the M2-1 tetramer as a flexible platform for RNA recognition, being the core domain the major binding site responsible for the first molecular contact with the nucleic acid molecules. Despite the well-described biological activity (11, 12) and structures of M2-1 (8, 9), and recent clarifications on the possible biochemical mechanism accounting for its antitermination activity (10), molecular details about biophysical-chemical and dynamic features of the interaction between the core domain of M2-1 and RNA still need to be introduced. In this report, an integrated approach with experimental and computational techniques was employed to obtain information at the molecular level on the binding of hRSV cdM2-1 to long RNAs (~80 bases). The fluorescence spectroscopy experiments at different conditions of temperatures, ionic strengths, and pH revealed a key role of the electrostatic interactions in the formation of the complex and that the van der Waals interactions

and hydrogen bonds in the bimolecular stabilization, while differential scanning calorimetry (DSC) measurements indicated a slight increase in thermal stability of cdM2-1 upon RNA binding. Nuclear magnetic resonance (NMR) highlighted structural features of the cdM2-1 and its complex with RNA. The chemical shift perturbation (CSP) analysis identified the α1–α2–α5–α6 helix bundle in the core domain as the main interaction region with RNA and, along with thermal susceptibility measurements of the amide hydrogen ($^{1}H_N$) chemical shifts, pointed out that this bundle helix may undergo local structural changes that extend to the α3–α4 hairpin. The solvent paramagnetic relaxation enhancement (sPRE) measurements showed that cdM2-1 is permanently protected from solvent exposure when complexed with the nucleic acid, while intensity changes in imino proton resonances suggest a protein-induced unfolding of partial formations of RNA base-pairing. ^{15}N relaxation data revealed that cdM2-1 dynamics undergo significant changes after interaction with RNA, presenting a decrease in the overall tumbling motion and locally an increase in the flexibility of the α2–α3 and α3–α4 loops of the protein. The structural models of the cdM2-1/RNA complex generated from docking calculations and evaluated by molecular dynamics (MD) simulations corroborated with experimental results and highlighted Lys150 and Arg151 as pivotal residues for the stabilization of the complex. The results presented herein contribute to understand the molecular behavior of the core domain of hRSV M2-1 in solution and upon binding to its biological partner, the RNA.

RESULTS

cdM2-1/RNA interaction investigated by fluorescence spectroscopy

The binding of the hRSV cdM2-1 to RNA was investigated using fluorescence quenching experiments at different temperatures (15, 25, and 35 °C), salt concentrations (0, 150, and 350 mM of NaCl), and pHs (5, 6.5, and 8). Fig. 1A shows that the fluorescence intensity of cdM2-1 is quenched with increments of RNA concentrations, suggesting that the microenvironment of the

153 fluorophores of the domain is affected by the presence of the nucleic acid. The changes in the
154 fluorescence quenching signal $(F_0 - F)/(F_0 - F_S)$ as a function of RNA concentrations were
155 analyzed using Eq. (1) at the temperature of 15, 25, and 35 °C (Fig. 1B). The binding isotherm
156 curves provided the values of the dissociation constant (K_d) and stoichiometry coefficient (n_{SC})
157 for the cdM2-1/RNA complex, which are shown in Table 1. The values of K_d determined for the
158 cdM2-1/RNA interaction are in the nano to micromolar range and the binding affinity revealed to
159 decrease with increasing temperature from 15 to 35 °C. The average value of n_{SC} (~2.25) at the
160 investigated temperatures indicates that at least two cdM2-1 binds to one RNA molecule.
161 Figure 1
162 The van't Hoff plot [Eq. (2)] ($\ln K_d$ & 1/T, Fig. 1B) was linear over the investigated
163 temperature range. The enthalpy change (ΔH) was –45 ± 10 kJ·mol^{-1} (Table 1), denoting an
164 enthalpically favorable exothermic binding reaction. The negative values of changes in Gibbs
165 free energy (ΔG) and entropy (ΔS) indicate that the interaction is spontaneous and entropically
166 unfavorable. The enthalpic term provides the major contribution to ΔG, suggesting that the
167 binding process is enthalpically driven. $\Delta H < 0$ and $\Delta S < 0$ indicate that van der Waals
168 interactions and hydrogen bonds are important non-covalent interactions responsible for the
169 stabilization of the cdM2-1/RNA complex (13).
170 Fig. 1C and 1D show the effect of ionic strength and pH at 25 °C. With an increase in the
171 salt concentration, the fluorescence quenching percentage rates [$(F_0/F - 1)\%$] decreased in each
172 titration step of RNA with cdM2-1 (Fig. 1B), indicating that the rising in ionic strength caused a
173 reduction in the quenching efficiency with a tendency to increase the RNA binding constant
174 (1.13 ± 0.11 and 1.70 ± 0.19 μM for n_{SC} = 2.25 at 0 and 350 mM NaCl, respectively). The
175 decrease in quenching efficiency is probably related to an electrostatic shielding of the protein
176 and/or nucleic acid. The decrease in pH from 8 to 5 induced an increase in $(F_0/F - 1)\%$ that

suggests an increase in the quenching efficiency of the cdM2-1 by RNA, which may be explained by the strengthening of the electrostatic interactions involved in the formation of the cdM2-1/RNA complex. Protonation of His147 and His168 at pH 5 may play a key role. Despite the increase in $(F_0/F-1)\%$, there was a decrease in the cdM2-1/RNA affinity (K_d = 0.90 ± 0.14 and 4.04 ± 0.41 μM for n_{SC} = 2.25 at pH 8 and 5, respectively) due to reduction in pH, which reinforces the role of electrostatic interactions in the formation of the complex while hydrogen bonds and van der Waals interactions play a key role in the binding stabilization. Circular dichroism (CD) experiments showed that there are no significant changes in the secondary structure elicited by the changes in pHs (data not shown). The variations in the fluorescence quenching percentage rates at different conditions of pH did not arise from secondary structure changes.

Table 1

Thermal stability of cdM2-1 and cdM2-1/RNA complex determined via DSC

The thermal denaturation analysis by differential scanning calorimetry (DSC) showed that cdM2-1 has an endothermic transition with a melting temperature (T_m) of 68.7 ± 0.4 °C and a calorimetry enthalpy change (ΔH_{cal}) of 174 ± 18 kJ.mol^{-1} (Fig. 2). Interestingly, Teixeira and collaborators (2017) showed a T_m value of ~55 °C for the full-length hRSV M2-1, which is at least 13 °C lower compared to that of its core domain (14). The van't Hoff enthalpy change (ΔH_{vH}) calculated from the integration of the cdM2-1 thermogram revealed a value of 341 ± 30 kJ.mol^{-1} that is different from ΔH_{cal}. This result indicated that the temperature-induced unfolding of cdM2-1 did not follow a two-state model. The thermal denaturation of the protein is irreversible, showing no transition in the cooling process. T_m and ΔH_{cal} values for the thermal unfolding of the cdM2-1/RNA complex were 69.4 ± 0.1 °C and 136 ± 10 kJ.mol^{-1} for a

protein/RNA molar ratio of 1:0.5, and 69.8 ± 0.1 °C and 103 ± 8 kJ.mol^{-1} for a molar ratio of 1:1, respectively (Fig. 2), indicating that the binding of cdM2-1 to RNA promoted a slight change in thermal stability of the protein.

Figure 2

Mapping cdM2-1 interaction with RNA

The analysis of chemical shift perturbation (CSP) by NMR spectroscopy was also used to characterize the cdM2-1/RNA binding (Fig. 3A). The CSP experiments presented a fast chemical exchange process between free and bound forms on the NMR chemical shift timescale (insert in Fig. 3A). The amino acid residues associated with resonances that showed chemical shift perturbation $\Delta\delta(^{1}H,^{15}N)$ higher than the average value plus standard deviation were (Fig. 3B and 3C): N-terminal (Leu74), α2-helix (Lys92, Gln93 side chain, Val97), α5-helix (His147, Lys150, Arg151), α5–α6 loop (Leu152), and α6-helix (Asp155, Val156). These residues of the dM2-1 promoted the binding to RNA either by direct interaction or by conformational rearrangements remote from the interaction interface. The resonances of residues Gly62, Ala63, Lys92, Gln93, Leu152, and Ala154 were broadened or even disappeared under the binding to the nucleic acid. The chemical shift changes of the amide group of [U-^{15}N] cdM2-1 under the binding to RNA revealed that the resonances of regions in the α1–α2–α5–α6 helix bundle presented the most significant $\Delta\delta(^{1}H,^{15}N)$ values, as seen by Blondot et al. (2012) (7). These regions denote a positive electrostatic potential (Fig. 3D) that is favorable for the interaction with a negative charge of the phosphate groups from the nucleic acid. In particular, Lys92, Lys150, and Arg151, which showed large chemical shift changes (Fig. 3B), are residues conserved among hRSV and hMPV M2-1. Recently, Gao et al. (2020) characterized these charged residues as essential for RNA interaction from the crystal structure of hRSV M2-1 complexed with 7mer RNA (15). It is worth mentioning that Lys150 and Arg151 were identified from mutagenesis assays as key residues for transcription enhancement *in vivo* by hRSV M2-1 (7).

227 Figure 3

228

229 **Solvent PRE for free and RNA-bound dgM2-1**

230 Solvent exposure measures of each cdM2-1 backbone amide group and its complex with RNA
231 were obtained from solvent paramagnetic relaxation enhancement (sPRE) experiments (Fig. 4).
232 The more negative is the sPRE effect ($I_{NH} \cdot [Gd]^{-1}$, Fig. 4A), the more exposed is the residue to
233 the solvent. For the domain in the absence of RNA, it is possible to observe that the greatest
234 sPRE effects occur in N and C-terminals of cdM2-1 and mainly in the N-terminal residues of α1-
235 helix, characterizing these regions as significantly exposed to the solvent. Significant sPRE
236 effects are also observed in virtually all C-terminal residues of the protein loop regions,
237 highlighting the α1–α2 loop (Fig. 4A). In general, the helical secondary structure regions have
238 the lowest sPRE effects, indicating that the amide groups of the residues in these regions are
239 protected from solvent exposure. An exception to this is α4-helix which exhibited an $i + 4$ pattern
240 of solvent-exposed residues (Ser122, Arg126, Thr130, Ile133, and Ser137) corresponding to a
241 nearly α-helix turn (Fig. 4B). This data revealed the presence of a solvent-accessible cleft
242 between α1–α2–α5–α6 helix bundle and α3–α4 hairpin (Fig. 4B).

243 For the cdM2-1/RNA complex, it can be seen a significant reduction of the sPRE effect
244 throughout all the residues of the domain, except the N and C-terminals (Fig. 4A). This result
245 indicates that the helical secondary structure regions, along with their loops, are permanently
246 protected from solvent exposure when cdM2-1 is complexed with long RNAs, while the
247 terminals of the domain remain exposed to the solvent as observed for free cdM2-1. Thus, N and
248 C-terminus of the core domain of hRSV M2-1 might be considered as an intrinsically negative
249 internal control for the performed experiments.

250 Figure 4

251

252 **Thermal susceptibility measurements of the amide hydrogen chemical shifts**

253 The thermal susceptibility of the amide hydrogen (^{1}HN) chemical shifts ($d\delta_{HN}/dT$) of
254 cdM2-1 was measured in the absence and presence of RNA. $d\delta_{HN}/dT$ is correlated to the
255 strength of the hydrogen bonds formed by ^{1}HN, being values smaller than –5.0 ppb·K^{-1}
256 interpreted as weak hydrogen bonds (more expandable) usually found in unstructured and loop
257 regions, and values higher than –5.0 ppb·K^{-1} as strong hydrogen bonds that are typical of
258 secondary structure elements (Fig. 5A, see Materials and Methods for details about the
259 interpretation of $d\delta_{HN}/dT$ values). In general, cdM2-1 residues in secondary structure elements
260 presented $d\delta_{HN}/dT$ values higher than –5.0 ppb·K^{-1}, excepting Ile146 in α5-helix (green sticks
261 in Fig. 5C). Ile146 showed an unusual $d\delta_{HN}/dT$ value (< –5.0 ppb·K^{-1}, Fig. 5A) for a residue in
262 secondary structure, which stands out as a possible point of break in the hydrogen bond network
263 of α5-helix. On the other hand, most values of $d\delta_{HN}/dT$ lower than –5.0 ppb·K^{-1} were reported
264 for residues in the N and C-terminus and loops between helices of the core domain. However,
265 Glu59, Ala63, Asp67, Thr69, Glu70, Glu71, and Ala73 in the N-terminal of the cdM2-1 showed
266 $d\delta_{HN}/dT$ > –5.0 ppb·K^{-1}, indicating that the amide group of these residues might form less
267 expandable hydrogen bonds, typically as it is observed for ^{1}HNs in secondary structures (16).
268 These residues are not structured in the N-terminal of cdM2-1 (7) and, interestingly, they are in
269 α-helix secondary structures in the full-length hRSV M2-1. This data revealed a helical
270 propensity of this region for the cdM2-1.

271 The binding of the cdM2-1 to long RNAs promoted significant changes of $d\delta_{HN}/dT$ values
272 for residues all over the domain (Fig. 5A). Values of $\Delta d\delta_{HN}/dT$ significantly different from
273 zero were taken into consideration (Fig. 5B). Most of these residues map the protein/RNA
274 interaction interface delimited by regions in the N-terminal, α1, α2, α5, and α6 (Fig. 5C), which
275 corroborates with the results of the CSP experiments (Fig. 3B). Leu74, Gly75, Val97, Ser100,

Asp109, Asp116, Asn141 side-chain, Ile146, Arg151, Asp155, Val156, Lys158, Lys159, and Ile161 showed an increase in $\left|d\delta_{HN}/dT\right|$ ($\Delta d\delta_{HN}/dT > 0$) which indicates that hydrogen bonds formed by ^{1}HN of these residues became more expandable, possibly weaker; while hydrogen bonds established by ^{1}HN of Lys92, Gln93 side-chain, Ser94, Ala95, Thr104, Val131, Thr145, His147, Leu157, and Thr160 turned into less expandable, possibly stronger hydrogen bonds ($\Delta d\delta_{HN}/dT < 0$). In particular, the decrease for Gln93 side chain could be explained by a direct hydrogen bond with RNA since it is exposed to the solvent in free cdM2-1 (Fig. 4A) (7); while the increase for Asn141 side chain could be due to a weakening of its hydrogen bond formed with Ser137 in α4–α5 loop (7), which possibly comes from a conformational rearrangement of the α3–α4 hairpin. In the cdM2-1/RNA binding interface, there are two clusters of residues with their ^{1}HN forming hydrogen bonds that became less and more expandable, being the first cluster composed of residues Lys92, Gln93 side-chain, Ser94, and Ala95 in the N-terminal of α2-helix and the second one, formed by Asp155, Val156, Lys158, Lys159, and Ile161 in the N-terminal of α6-helix. Altogether, these results indicate that α1–α2–α5–α6 helix bundle suffers local structural changes upon RNA binding and α3–α4 hairpin shows punctual conformational rearrangements remote from the binding region.

Figure 5

Evidence of cdM2-1-induced structural changes in RNA

The one-dimensional (1D) ^{1}H NMR spectra of protein/nucleic acid complexes reveal that most of the resonance signals from the nucleic acid and protein are overlapped, as it can be observed from amine/amide proton region (6–10 ppm) in Fig. 6. The only spectral region that is not disturbed by the protein resonances is the imino proton region from 11 to 15 ppm (17). The resonances of imino protons are directly involved in base-pairing, which is remarkable evidence of the secondary structure organization of the nucleic acid. Fig. 6 shows the imino proton signals

of the protein-free and bound forms of the RNA, as well as the spectrum of the free cdM2-1 as a negative control since proteins have no observable resonances at this spectral region. The 1D ^{1}H spectrum of the protein-free RNA presented resonance signals of imino protons at 11–15 ppm range (blue line, Fig. 6), which indicates a structural arrangement level due to partial formations of RNA base-pairing. On the other hand, the ^{1}H spectrum of the cdM2-1-bound RNA showed a significant decrease in the intensities of the imino proton signals (red line, Fig. 6), which suggests a protein-induced unfolding of the RNA in its secondary structure elements. No significant line broadening was observed in the amine/amide region (6–10 ppm, in Fig. 6) of the 1D ^{1}H spectra, which shows that the decrease in the intensities of the imino proton signals is not due to the larger size of the cdM2-1/RNA complex.

Figure 6

Backbone dynamics of the hRSV cdM2-1 and its complex with RNA

The ^{15}N relaxation data of the protein backbone by NMR on pico to nanosecond timescales revealed the molecular dynamics of the protein in the absence and presence of RNA. The results of the spin nuclear relaxation (^{15}N R_1, R_2, and hetNOE) experiments with the cdM2-1 and cdM2-1/RNA complex are shown in Table S1. In the presence of RNA, there was a significant change in the values of ^{15}N R_1, R_2, and hetNOE, highlighting the increase in the average value of ^{15}N R_2 from 11.32 ± 1.10 to 22.24 ± 2.07 s^{-1} upon RNA binding (Table S2), which confirmed the formation of the cdM2-1/RNA complex. The N and C-terminals presented low hetNOE values in the absence and presence of RNA, meaning that they are flexible in the free and bound state.

The reduced spectral density mapping, calculated from the relaxation parameters, provides insights into internal motions of a protein with no *a priori* assumptions about its diffusion model. The spectral density functions $J(0.87\omega_H)$, $J(\omega_N)$, and $J(0)$ are sensitive to overall and intramolecular motions on pico to nanosecond timescales. $J(0.87\omega_H)$ and $J(\omega_N)$ report fast

motions at picoseconds while $J(0)$ probes motions at nanoseconds, with possible contributions from the chemical or conformational exchange term. The values of $J(0.87\omega_H)$, $J(\omega_N)$, and $J(0)$ for the free and RNA-bound cdM2-1 were calculated from Eqs. (3) – (5) and are presented in Fig. 7A, 7B, and 7C. As with the relaxation parameters, the spectral density values of the cdM2-1 showed outstanding changes in the presence of RNA, except for the N and C-terminals. The highest and lowest values of $J(0.87\omega_H)$ and $J(0)$, respectively, for residues in the N and C-terminals indicate that these regions are undergoing fast internal motion on a picosecond timescale and are highly flexible even after the binding to long RNAs. The plots of $J(0.87\omega_H)$ & $J(0)$ (Fig. 7E and 7F) indicated that these two spectral densities were significantly correlated with a correlation coefficient of –0.934 and –0.931 for the cdM2-1 in the absence and presence of RNA, respectively. On the other hand, $J(\omega_N)$ & $J(0)$ plots (Fig. 7G and 7H) showed smaller correlation coefficients with values of 0.642 and –0.227 for the RNA-free and bound forms, respectively. It is noteworthy that $J(\omega_N)$ & $J(0)$ plot presented a negative correlation for the bimolecular complex, and positive for the free cdM2-1. The average values of $J(0.87\omega_H)$, $J(\omega_N)$, and $J(0)$ excluding the terminal residues for the free core domain were 5.2 ± 1.0 ps, 258 ± 10 ps, and 3.0 ± 0.3 ns, respectively; whereas the corresponding values for the RNA-bound cdM2-1 were 4.1 ± 1.6 ps, 198 ± 10 ps, and 6.1 ± 0.5 ns, respectively (Table S2).

Figure 7

R_2/R_1 ratios are proportional to the protein rotational correlation time (τ_C) and the values above the average may be involved in conformational exchange. The R_2/R_1 ratios calculated for the residues of cdM2-1 in the absence and presence of the nucleic acid demonstrated the complex formation, as the average R_2/R_1 value, excluding the N and C-terminal residues, increased from 8.0 ± 0.8 to 20.4 ± 2.2 ns upon addition of RNA (Fig. 7D). For the free cdM2-1, the R_2/R_1 values for Ile84, Ile87, and Asn88 in α1-helix and α1–α2 loop are above the average, suggesting that

these residues are in conformational exchange (Fig. 7D). After the formation of the cM2-1/RNA complex, the residues Ile87 and Asn88 in α1–α2 loop still presented R_2/R_1 values higher than the average while others such as Asp155, Val156, Asn163, and Thr164 in α6-helix had an increase in ratio values above the average, possibly due to chemical exchange (binding equilibrium). Interestingly, Asn138 in α4-helix with R_2/R_1 and $J(0)$ values above the average upon RNA binding reinforces the idea of α3–α4 hairpin conformational rearrangements remote from the interaction interface.

The extended Lipari-Szabo model-free formalism was used to analyze the relaxation parameters to obtain a hydrodynamic description of the protein informing the global diffusion anisotropy and the local motions of the cd-M2-1 (Table S3). A rotational correlation time (τ_C) for the protein tumbling of 8.53 ± 0.07 ns at 25 °C, compatible with the monomeric protein, was calculated using a fully anisotropic model for molecular rotational diffusion (see details in Material and Methods). For comparison with experimental data, the estimation of the τ_C value was calculated using HYDRONMR program (18) using the 20 NMR structures of the cdM2-1 (PDB access code 2L9J) (7), excluding amino acid residues from the flexible terminals, and the calculated τ_C was of 8.1 ± 0.3 ns at 25 °C. The values of $S^2 < 0.7$ for the N and C-terminal residues indicated significant thermal fluctuations and high degree of internal mobility for these protein regions (Fig. 8). Residues located in α1–α2 loop (Ile90), α3-helix (Asp110, Lys113, Leu114), α3–α4 loop (Glu119, Asn121, Ser122), α4–α5 loop (Asn141), and α5-helix (Ile146) showed significantly lower S^2 values than the helical-core average (0.91 ± 0.04, excluding terminal regions), which suggests a certain level of flexibility (Fig. 8A and 8C). The below-average S^2 along with $d\delta_{HN}/dT < -5.0$ ppb·K^{-1} for Ile146 in α5-helix reinforce the assumption that this residue is as a point of break of the cdM2-1 secondary structure. Conformational exchange contributions with $R_{ex} > 1.0$ Hz were observed for the residues Glu59, Ile84, Ile87,

Asn88, Ile90, Glu119, and Ala154 of the cdM2-1 (Fig. 8B and 8D). The model-free analysis of Lipari-Szabo for the RNA-bound cdM2-1 did not provide a satisfactory result.

Figure 8

Computational approach of the cdM2-1/RNA interaction

The structural models of the cdM2-1/RNA complex were generated using the 3dRPC web server (19). The structural restraints of the complex were defined from CSP analysis, defining residues as involved in the binding interface when $\Delta\delta > \Delta\delta_{average}$. Fig. 9A shows 10 structural models of the cdM2-1/RNA interaction determined by independent docking calculations, which correspond to the lowest energy structures from 10 complex models predicted by the 3dRPC-Score function (19). The 10 RNA molecules crowded around the α1–α2 and α5–α6 loops protruding like an "umbrella" over the α1, α2, α5, and α6 helices and the solvent accessible cleft formed between α1–α2–α5–α6 helix bundle and α3–α4 hairpin. Such behavior can be evidenced by the analysis of the mass density map of the protein-bound RNA (Fig. 9C) and by the change in absolute accessible surface area (ΔASA) of the RNA-bound cdM2-1 (Fig. 9D). The docking calculations also showed that the cdM2-1 preferably binds close to the secondary structure regions of the RNA molecules, where the nucleic acids presented a conformational organization level due to the base-pairing. This fact is exemplified by the cdM2-1/RNA complex model in Fig. 9B, where the α1–α2–α5–α6 helix bundle of the protein lies between two stem-loop secondary structures of the nucleic acid. The evaluation of the non-covalent interactions carried out with PLIP server (20) revealed that the residues Lys150 and Arg151 were involved in salt bridges and hydrogen bonds in nine and eight of 10 cdM2-1/RNA complex models, respectively (Table S4–S13). In general, Lys150 formed two salt bridges with phosphate groups of adjacent nucleotides, while Arg151 established hydrogen bonds with nitrogenous bases and ribose groups of the nucleic acids. These two residues also took place in π-cation interactions with the heterocyclic rings of nitrogenous

bases at least three of 10 structural models of the cdM2-1/RNA complex. Recently, the participation in salt bridges of Lys150 was also reported by Gao et al. (2020) for the crystal structure of the hRSV M2-1/7mer RNA complex (15).

Figure 9

The MD calculations were used to check the structural stability of cdM2-1/RNA complex models generated from the 3dRPC server. The number of hydrogen bonds, number of contacts, and RMSD analyzed from the 20 ns MD simulations for the complex models are presented as average values for five independent calculations in Fig. 10. RMSD values for non-hydrogen atoms of RNA and backbone atoms of the helical-core region of the domain (cdM2-$1_{75–171}$) reached stable levels after 2.5 and 1.0 ns, respectively, whereas for the entire cdM2-1 observed significant differences due to the flexible terminal regions contributions (Fig. 10A). The number of contacts between atoms of the cdM2-$1_{75–171}$ and RNA molecules for distances < 0.6 nm was almost stable, being approximately 5200 (Fig. 10B), indicating that the protein and nucleic acids remained interacting throughout the simulation time. Fig. 10C shows that the number of hydrogen bonds formed between the cdM2-$1_{75–171}$ and nucleic acids presented very stable along with all the MD simulations with an average value of around nine, being Lys150 and Arg151 responsible for at least one hydrogen bond each. Therefore, the molecular dynamics calculations proved the stability of the structural models of the cdM2-1/RNA complex over 20 ns simulations.

Figure 10

DISCUSSION

Here, it was presented the first evidence that cdM2-1 alone has an unfolding activity for long RNAs and therefore has an active role in the function of the full-length protein as a processivity and antitermination factor of the RdRp complex. The intensity of the imino proton resonances is a direct measure of the base-pairing that stabilize RNA secondary structures. The reduction in these intensities was explained by the destabilizing activity of cdM2-1 (Fig. 6), promoting an

425 unfolding of RNA secondary structures. This structural finding is different from that reported by
426 Molina et al. (2018), in which they observed no significant structural change from the 20mer
427 RNA band at 270 nm in presence of cdM2-1 via CD measurements (10). The evidence that the
428 cdM2-1 alone induced secondary structure unfolding in long nucleic acid molecules may be a
429 consequence of the molecular recognition of RNA base-pairing by this domain. Such logical
430 thinking sums with the outcome reported by Blondot et al. (2012), in which cdM2-1 presented a
431 higher affinity for the double-stranded RNA (12mer) than for its corresponding single-stranded
432 ones, highlighting the key role of base-pairing of the nucleic acid for the cdM2-1/RNA
433 interaction (7).

434 The molecular docking calculations revealed that the structural models of the cdM2-1/RNA
435 complex are supported by CSP and sPRE analysis (Fig. 3 and 4), which suggest that despite the
436 nucleic acids specifically bind to the distal portions of the α1, α2, α5, and α6 helices, the RNA
437 molecules may surround the entire helical-core region of the protein (Fig. 9A, 9C, and 9D). The
438 theoretical structural models of the complexes also corroborate with the evidence of the RNA
439 unfolding induced by cdM2-1 verified by NMR (Fig. 6), since the models indicate that the
440 domain preferably binds to the RNA secondary structure regions (Fig. 9B). Therefore, the
441 hypothesis is that the unfolding activity of long RNAs would be triggered by molecular
442 recognition of base-pairing.

443 The cdM2-1/RNA binding presents K_ds in the nano to micromolar range (Table 1), which
444 were similar to values reported by Blondot et al. (2012) and Molina et al. (2018) (7, 10). On the
445 other hand, the full-length M2-1 has K_ds of tenths of nanomolar that are at least 10-fold lower
446 than those recorded for the cdM2-1 (8, 10, 21). This indicates that although the core domain is
447 responsible for the first molecular contact with the nucleic acid molecules, there is another region
448 on M2-1 that provides an increase in the affinity with RNA. These regions may be the zinc finger
449 domains, which provide a type of binding specificity for adenine bases at specific positions in the

nucleotide sequence (9, 15), and/or the neighboring core domains in the tetrameric arrangement, promoting a positive cooperative binding to two RNA molecules of 13mer or longer per tetramer of the M2-1 (10). Interestingly, Gao et al. (2020) recently revealed from crystal structures and molecular dynamics of the hRSV M2-1/7mer RNA complex that the interactions of RNA with two separate domains of M2-1, the zinc finger domain and the core domain, can be independent of each other (15).

The temperature, ionic strength, and pH-dependence on the cdM2-1/RNA binding affinity suggests that electrostatic interactions drive the formation of the encounter complex, while van der Waals interactions and hydrogen bonds are responsible for the complex stabilization ($\Delta H < 0$ e $\Delta S < 0$) (13). An analysis of non-covalent interactions of the cdM2-1/RNA docking models shows the occurrence of hydrogen bonds, salt bridges, and π-cation interactions (Table S3–S12), which is in line with the interpretation of the thermodynamic parameters. These results are important to understand the M2-1/RNA interaction since they indicate that this process is fine-tuned by the molecular environment. In the cell, M2-1 is found in cytoplasmic inclusion bodies in which all the components of the viral polymerase concentrate and where the RNA synthesis occurs (22).

DSC data revealed that cdM2-1 is more thermal stable than the full-length M2-1. This behavior is not observed when the domains are intimately interacting, as tethered domains, in the multi-domain proteins (23). This indicates that the thermal stability of cdM2-1 is affected due to contacts in the full-length protein, which may promote a type of negative inter-domain coupling interaction, as defined by Bhaskar & Srinivasan (2011) (23). The results also pointed out a slight increase in thermal stability upon RNA binding, indicating a preferential binding to the native state of the core domain (24).

The α1–α2–α5–α6 helix bundle is the main region involved in the cdM2-1/RNA interaction, as determined by CSP. This helix bundle delimits a positive electrostatic potential surface on cdM2-1 (Fig. 3D) that contributes to the formation of the encounter complex, considering the

ionic strength and pH-dependence on binding affinity to RNA (Fig. 1C and 1D). The thermal susceptibility measurements of ^{1}HN chemical shifts suggest that the interaction with RNA promoted local structural rearrangements of the helix bundle, with residues in N-terminals of α2 and α6-helix showing stronger (less expandable) and weaker (more expandable) hydrogen bonds, respectively (Fig. 5C). Except for the N and C-terminals of cdM2-1, the sPRE experiments pointed out that the helical-core region is permanently protected from solvent exposure when complexed with the long RNA (Fig. 4A). This result, along with the analysis of CSP and thermal susceptibility of ^{1}HN chemical shifts, suggests that an RNA region anchors mainly in α1–α2 and α5–α6 loops while other non-interacting portions of its length protrude over the α1–α2–α5–α6 helix bundle, wrapping the helical region of cdM2-1.

The ^{15}N backbone dynamics data revealed that the N and C-terminals of cdM2-1 are equally flexible in the free and RNA-bound state, corroborating with the analyses of sPRE, CSP, and $d\delta_{HN}/dT$ which show that the terminal residues did not take part in the interaction with the nucleic acid and remained solvent-exposed. The free cdM2-1 helical-core has well-defined structures with few localized points of thermal flexibility in loops (Fig. 8), which may serve as hinges that confer plasticity to the domain. Residues in α1–α2 loop of free cdM2-1 undergo dynamics on micro to millisecond timescale (conformational exchange) that is important for the molecular recognition of RNA since this region participates directly in the interaction interface. There are significant changes in the dynamics of the RNA-bound cdM2-1, showing R_2/R_1 ratios that characterize a decrease in the overall tumbling motion when compared with the free domain and thus confirm the complexation. The values of $J(0)$ and R_2/R_1 ratio reveal that residues in α6-helix displayed dynamics on micro to millisecond timescale upon RNA binding, while the residues in α1–α2 loop remained in conformational exchange. It is worth pointing out that helices α1–α2 loop and α6-helix cdM2-1 are important for the binding to nucleic acids as it was

500 determined herein by CSP analysis and thermal susceptibility measurements and also reported by
501 Blondot et al. (2012) (7).

502 The best way to describe the dynamics of cdM2-1/RNA complex is through the reduced
503 spectral density mapping. Here, this analysis enriched the dynamical information, indicating that
504 the α2–α3 and α3–α4 loops of the protein became more flexible after the binding to RNA since
505 these undergo dynamics on picosecond timescale (fast internal motions). The mean $J(0)$ values
506 of the α2–α3 and α3–α4 loops decreased significantly with respect to the helical-core average for
507 the RNA-bound cdM2-1, while the mean $J(0.87\omega_H)$ values of these loops were slightly higher
508 than the overall average (Table S2). The inverse correlation between $J(0)$ and $J(0.87\omega_H)$ (Fig.
509 7E and 7F) has been widely reported (25–27) and it can be exemplified for Asn107 in α2–α3
510 loop and Glu119 in α3–α4 loop (Fig. 7A and 7C). The mean $J(\omega_N)$ values of the α2–α3 and α3–
511 α4 loops were significantly lower than the helical-core average for the free domain and, after the
512 interaction with RNA, these values approached the overall average (Table S2). Asp110, Lys113,
513 and Lys114 in α3-helix showed similar behavior to these loops (Fig. 7B). The direct and inverse
514 correlations observed from $J(\omega_N)$ & $J(0)$ plot for the protein in the absence and presence of
515 RNA (Fig. 7G and 7H), respectively, suggest that fast internal motions took place in α2–α3 and
516 α3–α4 loops and in α3-helix before complexation and after that, these fast dynamics increased.
517 This result corroborates with the analysis of $J(0)$ and $J(0.87\omega_H)$.

518 In conclusion, the present work describes that cdM2-1 alone has an unfolding activity for
519 long RNAs, which is probably triggered by the molecular recognition of the base-pairing. The
520 α1–α2–α5–α6 helix bundle characterizes as the major RNA binding region of cdM2-1, which
521 undergoes local conformational changes due to the interaction. Although the α3–α4 hairpin is not
522 directly involved in the binding to RNA, it presents conformational rearrangements remote from
523 the interaction interface and also an increase in picosecond motions of its α2–α3 and α3–α4

loops. The cdM2-1 as a whole is surrounded by the long RNAs and the most likely hypothesis is that the helix bundle preferentially binds to the base-pairing while its other regions protrude over the helical-core region of the domain. Therefore, it is revealed that the cdM2-1/RNA complex originates from a fine-tuning binding which likely contributes to the interaction aspects required to processivity and antitermination activity of the M2-1.

MATERIAL AND METHODS

Sample preparation

The core domain (residues 58–177) of hRSV M2-1 (cdM2-1) was expressed in *E. coli* BL21 (DE3) RIL cells with a pD441-NHT vector (ATUM, USA) in M9 minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source for the production of isotopically labeled protein, as described previously (28). The construct of the vector includes an N-terminal hexahistidine affinity tag (His_6-tag) and a TEV cleavage site. After the expression, the cell suspension was centrifuged and the pellet was resuspended in buffer A (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 1.0 mM β-mercaptoethanol, and 5% (v/v) glycerol). The cells were lysed by sonication and next the crude extract was centrifuged. The supernatant was loaded on a Ni-NTA column for affinity chromatography, pre-equilibrated with buffer A. The column was washed extensively with buffer A containing 10, 20, and 40 mM imidazole, and the protein was eluted with a 60 to 500 mM imidazole gradient. The eluted protein fractions were digested for removing of the His_6-tag using TEV protease at 20 °C for 14 hours. This step was performed via dialysis using a solution containing 20 mM Tris-HCl (pH 8.0), 1.0 mM DTT, and 0.5 mM EDTA. Next, the protein without His_6-tag was injected into a Superdex 75 10/300 GL (GE Healthcare) size exclusion column with buffer B (50 mM NaH_2PO_4/Na_2HPO_4 pH 6.5, 150 mM NaCl, and 1.0 mM DTT) used for fluorescence quenching, CD, DSC, and NMR experiments. In the fluorescence quenching and CD experiments, the protein sample was dialyzed against a citrate/phosphate and mono/dibasic phosphate buffer to reach the pH 5 and 8, respectively. For NMR measurements, it

was added 0.1% (w/v) NaN_3 and 7% (v/v) D_2O into phosphate buffer solution (buffer B). The RNA from Torula Yeast type VI (~80 bases) was purchased from Sigma-Aldrich. The stock solution of RNA was prepared at the same buffer solution as for protein, taking into account the pH control because of the solubilization of the nucleic acid. The concentrations of cdM2-1 and RNA were determined by absorbance at 280 and 260 nm using extinction coefficients of 5,960 $M^{-1} \cdot cm^{-1}$ (Expasy-ProtParam) (29) and 8,050 $M^{-1} \cdot cm^{-1} \cdot base^{-1}$ (ThermoFisher Scientific, DNA, and RNA Molecular Weights and Conversions), respectively.

Fluorescence quenching measurements

The fluorescence quenching experiments were performed using Fluorescent Spectrometer Lumina (Thermo Fisher Scientific, USA) equipped with a Peltier system for temperature control and a 10 mm optical path quartz cuvette. The excitation wavelength at 288 nm was set to promote the fluorescence emission of the protein. The emission spectra were reported in a 300–450 nm range with an increment of 1.0 nm, which was corrected for background fluorescence (from the buffer) and inner filter effects (30). Both excitation and emission bandwidths were set at 5 nm. Each point in the emission spectrum is the average of 10 accumulations. The titrations were performed adding small aliquots of the RNA stock solution (715 μM) to 2 mL of cdM2-1 at a constant concentration (5.5 μM). The titration experiments were reported at different temperatures (15, 25, and 35 °C) for determining the binding thermodynamic profile of the cdM2-1/RNA complex, as well as at different conditions of salt concentration (0, 150, and 350 mM of NaCl) and pH (5, 6.5, and 8). Measurements were performed in duplicate.

The fluorescence quenching data of the cdM2-1/RNA interaction were analyzed using the following equation (31):

$$\frac{F_0 - F}{F_0 - F_{sat}} = \frac{1}{2n_{SC}[P_T]}\left[\left(n_{SC}[P_T] + [L_T] + K_d\right) - \sqrt{\left(n_{SC}[P_T] + [L_T] + K_d\right)^2 - 4n_{SC}[P_T][L_T]}\right] \quad (1)$$

where F_0 is the fluorescence intensity in absence of the ligand (RNA), F the fluorescence in presence of RNA, F_{sat} the intensity of the bound protein saturated with RNA, K_d the dissociation constant, n_{SC} the stoichiometry coefficient of the cdM2-1/RNA complex corresponding to the number of protein molecules bound to one nucleic acid molecule, $[P_T]$ the total concentration of the protein, and $[L_T]$ the total concentration of the ligand. The values K_d and n_{SC} were determined from the fitting process to the experimental data by nonlinear least-squares optimization using Levenberg-Marquardt interactions.

Thermodynamic profile analysis

The driving forces responsible for the binding between biomolecules may include electrostatic interactions, hydrogen bonds, van der Waals interactions, and hydrophobic contacts (13). To elucidate the interactions involved in the cdM2-1/RNA complex, the thermodynamic parameters were calculated from the van't Hoff equation:

$$\frac{d}{d\frac{1}{T}}\ln\left(K_d\right)=\frac{\Delta H}{R} \quad (2)$$

where ΔH is the enthalpy change, R the universal gas constant, and K_d the dissociation constant at the correspondent temperature (T = 15, 25, and 35 °C). The ΔH value was obtained from the slope of the van't Hoff plot, with the respective values of Gibbs free energy changes (ΔG) determined from $\Delta G = RT\ln\left(K_d\right)$ and entropy change (ΔS) from $\Delta S = \left(\Delta H - \Delta G\right)/T$.

Circular dichroism measurements

The circular dichroism (CD) experiments were carried out using a Jasco 710 spectropolarimeter (Jasco, USA) equipped with a quartz cell of 0.5 mm optical path length. The far UV-CD spectra of 5.5 μM cdM2-1 were recorded at 25 °C and pHs 5, 6.5, and 8. The spectra were averaged over 20 scans in a 260–200 nm range with a resolution of 0.2 nm at the scan speed of 50 nm·min^{-1} and

598 1.0 nm spectral bandwidth. The CD signal taken as millidegrees (θ) was corrected for the
599 background contribution of the buffer and next expressed in terms of mean residues ellipticity
600 [Θ] in deg·cm^2·dmol^{-1} using $[\Theta] = \theta(m\,\mathrm{deg}) / (10[P] l n_R)$, where $[P]$ is the molar concentration
601 of the protein, l is the optical path length (cm), and n_R is the number of amino acid residues.
602
603 **Differential scanning calorimetry**
604 Differential scanning calorimetry (DSC) experiments were performed using an N-DSC III (TA
605 Instruments, USA) in the temperature range 10–90 °C at a heating and cooling scan rate of 1.0
606 °C/min. Both calorimetry cells were loaded with the buffer solution, equilibrated at 10 °C for 10
607 min and scanned repeatedly as described above until the baseline was reproducible. Next, the
608 sample cell was loaded with 100 μM of cdM2-1 plus 0, 50 and 100 μM of RNA, and scanned.
609 The thermograms of RNA in buffer solution were recorded as a control for the baseline. The
610 baseline corrections were obtained by subtracting the buffer (or RNA) scan from the
611 corresponding protein scan. Measurements were performed in duplicate. The calorimetry
612 enthalpy change (ΔH_{cal}) of the unfolding process of cdM2-1 was calculated from the area under
613 the thermogram curve. The van't Hoff enthalpy change (ΔH_{vH}) of the thermal denaturation
614 process was obtained from Eq. (2), replacing dissociation constant by unfolding constant (K_U).
615 The values of K_U were calculated by $f_U / (1 - f_U)$, where f_U is the fraction of unfolding protein
616 determined from an integral process of the denaturation thermogram.
617
618 **Chemical shift perturbation analysis**
619 The chemical shift perturbation (CSP) method was used to map the amino acid residues of
620 cdM2-1 involved in the binding to RNA. The two-dimensional (2D) ^{1}H–^{15}N HSQC experiments
621 were carried out in increasing amount of RNA from 0 to 115 μM which were added to the [U–
622 ^{15}N] cdM2-1 solution at constant concentration of 350 μM. The ^{15}N HSQC spectra were recorded

at 25 °C on NMR Bruker Avance III spectrometer of 14.1 T operating at a ^{1}H frequency of 600 MHz (Bruker BioSpin GmbH, Germany) equipped with cryogenically cooled Z-gradient probe. The data matrix of the spectra consisted of 1024* × 128* data points (were n* refers to complex points) with acquisition times 106.5 ms (t_{HN}) and 87.6 ms (t_N). A total of 16 scans per complex t_N increment were collected with a recycle delay of 1.25. The resonance assignment of the core domain of hRSV M2-1 was obtained from the repository BMRB (www.bmrb.wisc.edu) from the access code 17451 (28).

The ^{1}H-^{15}N HSQC spectra were processed using NMRPipe (32) and analyzed using CcpNMR Analysis (33). The chemical shift perturbation of the amino acid residues of the protein recorded in each titration step of RNA against [U–^{15}N] cdM2-1 solution was normalized using $\Delta\delta = \left(\Delta\delta_{HN}^2 + \Delta\delta_N^2/25\right)^{1/2}$ (34), where $\Delta\delta_{HN}$ and $\Delta\delta_N$ denote the chemical shift difference of ^{1}H and ^{15}N, respectively, recorded in absence and presence of RNA. The mean $\Delta\delta$ of the amino acid residues as well as its standard deviation of CSP were used as cutoff value to identify the residues of the protein involved in the cdM2-1/RNA interaction.

Amide hydrogen chemical shift temperature coefficient

The amide hydrogen ($^{1}H_N$) chemical shift temperature coefficient of cdM2-1 was determined by recording a series of 2D ^{1}H–^{15}H HSQC spectra at 15, 20, 25, 30, and 35 °C in the absence and presence of 115 μM RNA, using a Bruker Avance III NMR spectrometer of 14.1 T operating at a ^{1}H frequency of 600 MHz (Bruker BioSpin GmbH, Germany). All spectra were referenced to the water signal for each temperature, next processed using NMRPipe (32), and analyzed using CcpNMR Analysis (33). The chemical shift values (δ_{HN}) of all residues at different temperatures were plotted as a function of temperature. The resulting slope ($d\delta_{HN}/dT$) of every curve was plotted for each residue. For evaluation of the $d\delta_{HN}/dT$ values, it was used a straightforward

interpretation compilation as published by Morando et al. (2019) (35). In principle, $^{1}H_{N}$ chemical shift temperature coefficient reports on the thermal susceptibility of the H_N–C′ amide hydrogen bonds, since $d\delta_{HN}/dT$ is correlated to the length of the hydrogen bond ($r_{HNC'}$) and hydrogen bond J coupling $^{3h}J_{NC'}$ (16). On the one hand, residues with $d\delta_{HN}/dT$ < –5.0 ppb·K^{-1} form more expandable hydrogen bonds, which may be interpreted as weaker since they present smaller hydrogen bond J coupling $^{3h}J_{NC'}$. For this reason, an amide group in the protein structure with $d\delta_{HN}/dT$ < –5.0 ppb·K^{-1} may be considered a weak point of a secondary structure or, when this amide in a loop or intrinsically disordered region (IDR), more exposed to a hydrogen bond with water. On the other hand, residues with $d\delta_{HN}/dT$ > –5.0 ppb·K^{-1} make less expandable hydrogen bonds, which may be interpreted as stronger since they present larger hydrogen bond J coupling $^{3h}J_{NC'}$. In this sense, an amide group with $d\delta_{HN}/dT$ > –5.0 ppb·K^{-1} is involved in secondary structures of hydrogen bonds or, when this amide in a loop or IDR, may tend to form intramolecular hydrogen bonds and consequently to increase order (16, 35).

Paramagnetic relaxation enhancement experiments

To monitor the solvent accessibility of residues of cdM2-1 (350 μM) in the absence and presence of 115 μM RNA, it was used the gadolinium-based paramagnetic relaxation agent Gd-diethylenetriamine pentaacetic acid-bismathylamide (Gd-DTPA) (GE Life Science, United Kingdom). The $^{1}H_{N}/^{15}N$ amide cross-peak intensities of the protein were determined by recording a series of 2D ^{1}H–^{15}H HSQC spectra at different concentration of Gd-DTPA: 0, 1.0, 2.0, 3.0, and 4.0 mM at 25 ºC using a Bruker Avance III NMR spectrometer of 14.1 T operating at a ^{1}H frequency of 600 MHz (Bruker BioSpin GmbH, Germany). All spectra were processed using NMRPipe (32) and analyzed using CcpNMR Analysis (33). The intensities (I_{NH}) of each $^{1}H_{N}/^{15}N$ cross-peak of the residues were plotted as a function of Gd-DTPA concentration and the

slope of the adjusted straight line ($I_{NH} \cdot [Gd]^{-1}$) were plotted for each residue. To calculate an accurate slope $I_{NH} \cdot [Gd]^{-1}$, the points where the intensities are close to zero (lack of linearity) were not taken into account.

Imino proton resonances of RNA

Unidimensional ^{1}H excitation sculpting spectra were collected at 25 °C on Bruker Avance III NMR spectrometer of 14.1 T operating at a ^{1}H frequency of 600 MHz (Bruker BioSpin GmbH, Germany) using 180° water-selective pulse of 2 ms for the solvent suppression. The free induction decays (FIDs) were recorded with 32,768 data points using a spectral width of 21,000 Hz (35 ppm), a relaxation delay of 1.25 s, and an acquisition time of 0.78 s. The experiment was performed for the free protein at 350 μM and the free nucleic acid at 115 μM and after that, for the cdM2-1/RNA complex maintaining the same concentrations as before. The spectral region from 10 to 18 ppm was used for analyzing structural changes of RNA induced by cdM2-1, probing the imino protons of the nucleic acid involved in the base pair formation (17).

NMR relaxation experiments

The backbone dynamics of the [U–^{15}N] cdM2-1 at 350 μM and the core domain complexed with 115 μM of RNA were investigated from ^{15}N nuclear spin relaxation experiments (36) at 25 °C by using Bruker Avance III NMR spectrometer of 14.1 T operating at a ^{1}H frequency of 600 MHz (Bruker BioSpin GmbH, Germany). The R_1 experiments were collected using 14 delay times of inversion recovery of 54, 104, 204, 304, 404 (twice), 604, 804, 904 (twice), 1204, 1504 (twice), and 1804 ms. The $R_{1\rho}$ experiments were performed according to Korzhnev et al. (2002) (37) with 14 spin-lock periods of 10, 20, 30 (twice), 40, 50 (twice), 60, 70, 80, 90 (twice), 100, and 110 ms, and ^{15}N spin-lock field strengths of 2.0 kHz. A recycle delay of 3.0 s was used for the R_1

and $R_{1\rho}$ experiments. The values of R_1 and $R_{1\rho}$ were determined from non-linear least-square fittings of the intensities using two-parameter mono-exponential equations. R_2 values were determined from the measured R_1 and $R_{1\rho}$ rates (38). The reported errors for the rates were calculated from the estimated uncertainties for the relaxation delay duplicates. $\{^1H\}-^{15}N$ steady-state heteronuclear nuclear Overhauser effects (NOE) were measured from pairs of interleaved spectra recorded with (NOE) and without (control) H^N proton saturation during a recycle delay of 12 s. The $\{^1H\}-^{15}N$ hetNOE values were calculated from resonance intensity ratios obtained from the NOE and control spectra, with uncertainties estimated from the background noise of the spectra. The data were processed using NMRPipe (32) and analyzed using CcpNMR Analysis (33).

Reduced spectral density mapping approach

The reduced spectral density mapping (rSDM) is a simplified approach for analysis of nuclear spin relaxation data (R_1, R_2 and hetNOE) developed by Farrow et al. (1995) (39), Ishima & Nagayama (1995) (40), and Lefèrve et al. (1996) (41). The rSDM maps the spectral density for determining the accurate values of the spectral density function at three frequencies: $J(0)$ at the zero frequency, $J(\omega_N)$ at the nitrogen frequency, and $\langle J(\omega_H)\rangle$ at an effective proton frequency. The approach exploits the assumption that at higher frequencies, the spectral density terms contributing to the relaxation processes are equal in magnitude: $J(\omega_H) \approx J(\omega_H \pm \omega_N)$; and is replaced by a single equivalent term $\langle J(\omega_H)\rangle$, which corresponds to $J(0{,}87\omega_H)$. This assumption is based on the premise that at higher frequencies in comparison with $J(0)$ and $J(\omega_N)$, spectral densities $J(\omega_H)$ e $J(\omega_H \pm \omega_N)$ show minimal variation. The spectral density functions at three frequencies, as derived through reduced spectral density mapping, are given by the following equations:

$$J(0.87\omega_H) = \frac{1}{5d^2} 4\sigma_{NOE} \quad (3)$$

$$J(\omega_N) = \frac{1}{3d^2 + 4c^2}\left(4R_1 - 5\sigma_{NOE}\right) \quad (4)$$

$$J(0) = \frac{1}{3d^2 + 4c^2}\left(6R_2 - 3R_1 - 2.72\sigma_{NOE}\right) \quad (5)$$

where the cross-relaxation rate is defined by $\sigma_{NOE} = (hetNOE - 1)R_1 \gamma_N / \gamma_H$, $c = (\omega_N \Delta\sigma)/\sqrt{3}$, $d = \left(\frac{\mu_0 \hbar \gamma_N \gamma_H}{4\pi}\right)\frac{1}{r_{NH}^3}$, μ_0 is the permeability of the vacuum, $\hbar$ is reduced Planck constant, γ_N and γ_H are the gyromagnetic ratios of ^{15}N and ^{1}H, respectively, r_{NH} is the bond length, ω_N is the Larmor frequency of the ^{15}N nucleus, and $\Delta\sigma$ is the ^{15}N chemical shift anisotropy in ppm.

Model free analysis

The relaxation parameters were fitted according to the extended model-free formalism of Lipari-Szabo for obtaining the intramolecular dynamics (42). TENSOR2 program (43) was employed to define a motional model for hRSV cdM2-1 using an anisotropic tensor with a fully asymmetry diffusion model, (D_X = 1.81 ± 0.03, D_Y = 1.95 ± 0.02, D_Z = 2.11 ± 0.03) × 10^7 s^{-1}. The calculations of the rotational correlation time (τ_C) were performed from the relaxation data for the amino acid residues with values of R_2/R_1 ratios within one standard deviation relative to the calculated average and with values of {^{1}H}–^{15}N NOE heteronuclear higher than 0.65. The anisotropic diffusion tensor also was used for calculating the internal molecular dynamics parameters, such as order parameter (S^2) and conformational exchange rate (R_{ex}), which reflects movements of the NH bond in pico to nanosecond and micro to millisecond timescales, respectively.

Charge, protonation state, and electrostatic potential calculation

741 The electrostatic potential calculations were performed by the APBS software (44) using charge
742 values and protonation states obtained from the PDB2PQR server (version 2.1.1) (45) along with
743 the PROPKA program (version 3.0) (46). The physical-chemical parameters used for the
744 calculations were 150 mM NaCl, pH 6.5, and 25 °C. The electrostatic potential surface of the
745 core domain of hRSV M2-1 was displayed using PyMOL (47).
746
747 **Molecular modeling and molecular docking**
748 The primary sequences of 10 RNAs were generated from the Random Sequence Generator
749 (RSG) webserver (www.molbiotools.com) setting the length of sequence equal to 40 and AU
750 content at 80%. All RNA sequences generate by the RSG server are presented in Table S14.
751 These random RNA sequences were submitted to Direct Coupling Analysis (DCA) server
752 (www.biophy.hust.edu.cn) which is a statistical inference framework used to infer direct co-
753 evolutionary couplings among nucleotide pairs in multiple sequence alignments, which aims at
754 disentangling direct from indirect correlations. The information provided by the DCA server was
755 subsequently used in the 2dRNA web server (www.biophy.hust.edu.cn) which is a secondary
756 structure prediction method of RNA based on DCA data. The parameters of the 2dRNA server
757 were set as default. Next, the predicted secondary structure information was used in the 3dRNA
758 web server (version 2.0) (19) which is an automatic molecular modeling method for building the
759 three-dimensional structure of RNA. The assembly calculations performed by the 3dRNA server
760 were followed by an optimization step using all the parameters as default.
761 The molecular docking calculations for predicting the molecular model of the complex
762 formed between the random RNA sequences (Table S14) and cdM2-1 were performed by using
763 the 3dRPC web server (version 2.0) (48). The three-dimensional structure of the hRSV M2-1
764 core domain was downloaded from the Protein Data Bank, access code 2L9J (7), and 3D
765 structures of 10 RNAs were obtained from 3dRNA web server (version 2.0) (19). A total of 10
766 complex models were predicted using the 3dRPC-Score function. The amino acid residues of

cdM2-1 involved in the RNA binding interface, which was determined from the CSP method by NMR experiments, were specified as a constraint set in the advanced setting of the 3dRPC server. Following docking, the lowest energy structural models of 10 cdM2-1/RNA complexes were analyzed by VMD software (49) using the VolMap tool for evaluating of the mass density map of collocated RNA (1.0 Å of resolution; 1.5 Å × radius of atom size; 10 frames combined using average; and isosurfaces with isovalue of 0.5) and by PLIP webserver (20) for characterizing the protein/nucleic acid non-covalent interactions, such as hydrogen bond, π-cation interaction, π-stacking, and salt bridge. Structural conformation of the constructed models was displayed using PyMOL (47) and VMD (49).

Molecular dynamics simulation

The molecular dynamics (MD) simulations were carried out with the GROMACS program (version 5.0.7) (50). The molecular systems were modeled by using the AMBER99SB-IDLN protein and AMBER94 nucleic acid force field (51), and the TIP3P water model (52). The three-dimensional structures of 5 random-selected cdM2-1/RNA complexes calculated from molecular docking (3dRPC server) were used in the MD simulations. These structures were placed in the center of 89–135 Å cubic boxes filled with TIP3P water molecules and Na^+/Cl^- ions ([NaCl] = 150 mM). The protonation states of charged residues were set considering a pH 6.5 from the PROPKA results (46). All simulations were performed in NPT ensemble using periodic boundary conditions and keeping the system at 298 K (Nose-Hoover thermostat, τ_T = 2.0 ps) and 1.0 bar (Parrinello-Rahman barostat, τ_P = 2.0 ps and compressibility = 4.5×10^{-5} bar^{-1}). Lennard-Jones and Coulomb potentials were applied using a cutoff distance of 12 Å. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm. The covalent bonds involving hydrogen atoms were constrained to their equilibrium distance. A conjugate gradient minimization algorithm was utilized to relax the superposition of atoms

generated in the box construction process. The energy minimizations were performed with steepest descent integrator and conjugate gradient algorithm, using 500 $kJ \cdot mol^{-1} \cdot nm^{-1}$ as maximum force criterion. 100 thousand steps of molecular dynamics were performed for each NVT and NPT equilibration, applying force constants of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ to all heavy atoms of the cdM2-1/RNA complexes. Lastly, MD simulations of 20 ns were accomplished for data acquisition. Following the simulations, the trajectories were concatenated and analyzed according to the number of hydrogen bonds (cutoff distance = 3.5 Å and maximum angle = 30°), number of contacts (< 0.6 nm), and root-mean-square deviation (RMSD) of non-hydrogen atoms for the nucleic acid and backbone atoms for cdM2-1 (entire protein and helical-core region). These parameters (hydrogen bonds, number of contact, RMSD) from 5 MD simulations were presented as averages.

Accessible surface area calculations

The accessible surface areas (ASA) of free cdM2-1 and its docked complex with RNAs were calculated using the NACCESS program (53). The structural models of the cdM2-1/RNA complexes were obtained from the molecular docking calculations (3dRPC server). Changes in absolute ASA for residue i were calculated using $\Delta ASA^i = ASA^i_{free} - ASA^i_{complexed}$, where $ASA^i_{uncomplexed}$ and $ASA^i_{complexed}$ is the absolute accessible surface area for free and RNA-bound cdM2-1 residues, respectively. The values of ΔASA were averaged and next the average value was normalized to a range between 0 and 1.0 units for coloring the cdM2-1 structure.

ACKNOWLEDGMENTS

The author I.P.C. gratefully acknowledges the financial support by a postdoctoral fellowship from FAPERJ (202.279/2018) and the PROPe UNESP. G.C.G., V.B.M. and F.P.S. thank to the FAPESP scholarship (2019/08739-1, 2018/08900-4). The authors thank Prof. Dr. Marcelo de

Freitas Lima for the access to the Fluorescent Spectrometer Lumina located in the Bio-organic Environmental Laboratory of the Department of Chemistry and Environmental Sciences, and Prof. Dr. Alexandre Suma de Araújo for the access to the Calix cluster (FAPESP 2010/18169-3) for performing the molecular dynamics simulations. The authors also recognize GridUNESP for the availability of the software package for molecular dynamics simulations. The authors thank for the access to the NMR laboratory of Multiuser Center for Biomolecular Innovation (FAPESP 2009/53989-4) in IBILCE/UNESP. The authors thank NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS).

FUNDING

Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ, Brazil: Grant 202.279/2018, 210.361/2015, 239.229/2018, and 204.432/2014. Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil: 309564/2017-4. Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, Brazil: Grant 2019/08739-1, 2018/08900-4, 2010/18169-3, 2009/53989-4).

CONFLICT OF INTERESTS

The authors declare no conflict of interests exists.

REFERENCES

1. Collins PL, Karron RA. 2013. Respiratory syncytial virus and metapneumovirus, p. 1086–1123. *In* Knipe, DM, Howley, PM (eds.), Fields Virology: Sixth editionSixth Edit. Wolters Kluwer Health Adis (ESP).
2. Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, Auinger P, Griffin

842 MR, Poehling KA, Erdman D, Grijalva CG, Zhu Y, Szilagyi P. 2009. The Burden of respiratory
843 syncytial virus infection in young children. N Engl J Med 360:588–598.
844 3. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. 2005. Respiratory syncytial virus
845 infection in elderly and high-risk adults. N Engl J Med 352:1749–1759.
846 4. Bermingham A, Collins PL. 1999. The M2-2 protein of human respiratory syncytial virus is a
847 regulatory factor involved in the balance between RNA replication and transcription. Proc Natl
848 Acad Sci U S A 96:11259–11264.
849 5. Collins PL, Hill MG, Cristina J, Grosfeld H. 1996. Transcription elongation factor of respiratory
850 syncytial virus, a nonsegmented negative-strand RNA virus. Proc Natl Acad Sci U S A 93:81–85.
851 6. Hardy RW, Wertz GW. 1998. The Product of the Respiratory Syncytial Virus M2 Gene ORF1
852 Enhances Readthrough of Intergenic Junctions during Viral Transcription. J Virol 72:520–526.
853 7. Blondot M-L, Dubosclard V, Fix J, Lassoued S, Aumont-Nicaise M, Bontems F, Eléouët J-F,
854 Sizun C. 2012. Structure and Functional Analysis of the RNA- and Viral Phosphoprotein-Binding
855 Domain of Respiratory Syncytial Virus M2-1 Protein. PLoS Pathog 8:e1002734.
856 8. Tanner SJ, Ariza A, Richard CA, Kyle HF, Dods RL, Blondot ML, Wu W, Trincão J, Trinh CH,
857 Hiscox JA, Carroll MW, Silman NJ, Eléouët JF, Edwards TA, Barr JN. 2014. Crystal structure of
858 the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and
859 implications of its phosphorylation. Proc Natl Acad Sci U S A 111:1580–1585.
860 9. Leyrat C, Renner M, Harlos K, Huiskonen JT, Grimes JM. 2014. Drastic changes in
861 conformational dynamics of the antiterminator M2-1 regulate transcription efficiency in
862 pneumovirinae. Elife 2014.
863 10. Molina IG, Esperante SA, Marino-Buslje C, Chemes LB, de Prat-Gay G. 2018. Cooperative RNA
864 Recognition by a Viral Transcription Antiterminator. J Mol Biol 430:777–792.
865 11. Fearns R, Collins PL. 1999. Role of the M2-1 Transcription Antitermination Protein of
866 Respiratory Syncytial Virus in Sequential Transcription. J Virol 73:5852–5864.
867 12. Hardy RW, Wertz GW. 2000. The Cys3-His1 Motif of the Respiratory Syncytial Virus M2-1

868 Protein Is Essential for Protein Function. J Virol 74:5880–5885.
869 13. Ross PD, Subramanian S. 1981. Thermodynamics of Protein Association Reactions: Forces
870 Contributing to Stability. Biochemistry 20:3096–3102.
871 14. Teixeira TSP, Caruso ÍP, Lopes BRP, Regasini LO, Toledo KA de, Fossey MA, Souza FP de.
872 2017. Biophysical characterization of the interaction between M2-1 protein of hRSV and
873 quercetin. Int J Biol Macromol 95:63–71.
874 15. Gao Y, Cao D, Pawnikar S, John KP, Ahn HM, Hill S, Ha JM, Parikh P, Ogilvie C, Swain A,
875 Yang A, Bell A, Salazar A, Miao Y, Liang B. 2020. Structure of the Human Respiratory
876 Syncytial Virus M2-1 Protein in Complex with a Short Positive-Sense Gene-End RNA. Structure
877 28:979-990.e4.
878 16. Hong J, Jing Q, Yao L. 2013. The protein amide 1HN chemical shift temperature coefficient
879 reflects thermal expansion of the N-H···O=C hydrogen bond. J Biomol NMR 55:71–78.
880 17. Kyogoku Y, Kojima C, Jeon Lee S, Tochio H, Suzuki N, Matsuo H, Shirakawa M. 1995. Induced
881 structural changes in protein-DNA complexes. Methods Enzymol 261:524–541.
882 18. De La Torre JG, Huertas ML, Carrasco B. 2000. HYDRONMR: Prediction of NMR Relaxation
883 of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. J Magn
884 Reson 147:138–146.
885 19. Jian W, Mao K, Zhao Y, Chen Z, Jianjin X, Yi Z, Xiao Y. 2017. Optimization of RNA 3D
886 structure prediction using evolutionary restraints of nucleotide-nucleotide interactions from direct
887 coupling analysis. Nucleic Acids Res 45:6299–6309.
888 20. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. 2015. PLIP: Fully automated
889 protein-ligand interaction profiler. Nucleic Acids Res 43:W443–W447.
890 21. Cuesta I, Geng X, Asenjo A, Villanueva N. 2000. Structural Phosphoprotein M2-1 of the Human
891 Respiratory Syncytial Virus Is an RNA Binding Protein. J Virol 74:9858–9867.
892 22. Richard CA, Rincheval V, Lassoued S, Fix J, Cardone C, Esneau C, Nekhai S, Galloux M,
893 Rameix-Welti MA, Sizun C, Eléouët JF. 2018. RSV hijacks cellular protein phosphatase 1 to

regulate M2-1 phosphorylation and viral transcription. PLoS Pathog 14.

23. Bhaskara RM, Srinivasan N. 2011. Stability of domain structures in multi-domain proteins. Sci Rep 1.

24. Cooper A, Nutley MA, Wadood A. 2000. Differential scanning microcalorimetry, p. 287–318. *In* Harding, SE, Chowdhry, BZ (eds.), Protein-Ligand Interactions: Hydrodynamics and Calorimetry. Oxford University Press, Oxford New York.

25. Dayie KT, Wagner G, Lefèvre J-F. 1996. Theory and practice of nuclear spin relaxation in proteins. Annu Rev Phys Chem 47:243–282.

26. Bracken C, Carr PA, Cavanagh J, Palmer AG. 1999. Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: Implications for the entropy of association with DNA. J Mol Biol 285:2133–2146.

27. Berlow RB, Martinez-Yamout MA, Dyson HJ, Wright PE. 2019. Role of Backbone Dynamics in Modulating the Interactions of Disordered Ligands with the TAZ1 Domain of the CREB-Binding Protein. Biochemistry 58:1354–1362.

28. Dubosclard V, Blondot ML, Eléouët JF, Bontems F, Sizun C. 2011. 1H, 13C, and 15N resonance assignment of the central domain of hRSV transcription antitermination factor M2-1. Biomol NMR Assign 5:237–239.

29. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server, p. 571–607. *In* Walker, MJ (ed.), The Proteomics Protocols Handbook. Humana Press.

30. Lakowicz JR. 2006. Instrumentation for Fluorescence Spectroscopy, p. 27–61. *In* Principles of Fluorescence Spectroscopy. Springer US.

31. Beckett D. 2011. Measurement and analysis of equilibrium binding titrations: A beginner's guide, p. 1–16. *In* Johnson, ML, Holt, JM, Ackers, GK (eds.), Methods in Enzymology. Academic Press Inc.

32. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: A

multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293.

33. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. 2005. The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct Funct Genet 59:687–696.

34. Grzesiek S, Stahl SJ, Wingfield PT, Bax A. 1996. The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. Biochemistry 35:10256–10261.

35. Morando MA, Barbosa GM, Cruz-Oliveira C, Da Poian AT, Almeida FCL. 2019. Dynamics of Zika Virus Capsid Protein in Solution: The Properties and Exposure of the Hydrophobic Cleft Are Controlled by the α-Helix 1 Sequence. Biochemistry 58.

36. Farrow NA, Muhandiram R, Pascal SM, Kay LE, Singer AU, Forman-Kay JD, Kay CM, Gish G, Pawson T, Shoelson SE. 1994. Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. Biochemistry 33:5984–6003.

37. Korzhnev DM, Skrynnikov NR, Millet O, Torchia DA, Kay LE. 2002. An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates. J Am Chem Soc 124:10743–10753.

38. Tjandra N, Wingfield P, Stahl S, Bax A. 1996. Anisotropic rotational diffusion of perdeuterated HIV protease from 15N NMR relaxation measurements at two magnetic fields. J Biomol NMR 8:273–284.

39. Farrow NA, Zhang O, Szabo A, Torchia DA, Kay LE. 1995. Spectral density function mapping using 15N relaxation data exclusively. J Biomol NMR 6:153–162.

40. Ishima R, Nagayama K. 1995. Protein Backbone Dynamics Revealed by Quasi Spectral Density Function Analysis of Amide N-15 Nuclei. Biochemistry 34:3162–3171.

41. Lefèvre JF, Dayie KT, Peng JW, Wagner G. 1996. Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. Biochemistry 35:2674–2686.

946 42. Lipari G, Szabo A. 1982. Model-Free Approach to the Interpretation of Nuclear Magnetic
947 Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. J Am Chem Soc
948 104:4559–4570.
949 43. Dosset P, Hus JC, Blackledge M, Marion D. 2000. Efficient analysis of macromolecular
950 rotational diffusion from heteronuclear relaxation data. J Biomol NMR 16:23–28.
951 44. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems:
952 Application to microtubules and the ribosome. Proc Natl Acad Sci U S A 98:10037–10041.
953 45. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA. 2007. PDB2PQR:
954 expanding and upgrading automated preparation of biomolecular structures for molecular
955 simulations. Nucleic Acids Res 35:W522-5.
956 46. Olsson MHM, SØndergaard CR, Rostkowski M, Jensen JH. 2011. PROPKA3: Consistent
957 treatment of internal and surface residues in empirical p K a predictions. J Chem Theory Comput
958 7:525–537.
959 47. Delano WL. 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA,
960 USA. San Carlos, CA, USA.
961 48. Li H, Huang Y, Xiao Y. 2017. A pair-conformation-dependent scoring function for evaluating 3D
962 RNA-protein complex structures. PLoS One 12:e0174662.
963 49. Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. J Mol Graph
964 14:33–38.
965 50. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindah E. 2015. Gromacs: High
966 performance molecular simulations through multi-level parallelism from laptops to
967 supercomputers. SoftwareX 1–2:19–25.
968 51. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010.
969 Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins Struct
970 Funct Bioinforma 78:1950–1958.
971 52. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple

potential functions for simulating liquid water. J Chem Phys 79:926–935.

53. Hubbard SJ, Thornton JM. 1993. NACCESS: computer program. Department of Biochemistry and Molecular Biology, University College, London.

FIGURE LEGENDS

Figure 1. Analysis of the fluorescence quenching data of the cdM2-1/RNA interaction at different conditions of temperature, ionic strength, and pH. (A) Emission spectra of the cdM2-1 in absence and presence of RNA concentration increments (pH 6.5, T = 25 °C, λex = 288 nm). [cdM2-1] = 5.5 μM; [RNA] = 0–16.8 μM. The fluorescence quenching spectra at 15 and 35 °C showed a similar profile to 25 °C (data not shown). (B) The analysis of the fluorescence data of the cdM2-1/RNA interaction. Changes of the fluorescence quenching signal of cdM2-1 as a function of RNA concentrations in 50 mM phosphate buffer (pH 6.5) containing 150 mM NaCl and 1.0 mM DTT at 15, 25, and 35 °C. The insert denotes the van't Hoff plot used to determine the enthalpy change value in the formation of the cdM2-1/RNA complex. (C, D) Fluorescence quenching percentage rates as a function of RNA concentrations at different conditions of salt concentration (0, 150, and 350 mM NaCl) and pH (5, 6.5, and 8) for the temperature of 25 °C.

Figure 2. Thermal stability of the free and RNA-bound cdM2-1. DSC thermograms of the protein and its complex with RNA collected in 50 mM phosphate buffer (pH 6.5) containing 150 mM NaCl and 1.0 mM DTT using scan rate of 1.0 °C/min. The dotted lines denote the melting temperature (T_m) for cdM2-1 in absence and presence of RNA (100 μM). The concentration of the protein was of 100 μM and nucleic acid of 0 (black square), 50 (red triangle), and 100 μM (green circle).

Figure 3. cdM2-1/RNA binding investigated by NMR spectroscopy. (A) Two-dimensional $^{1}H–^{15}N$ HSQC spectra of the free (black) and RNA-bound [U-^{15}N]cdM2-1 (red) collected by using NMR spectrometer of 14.1 T (^{1}H frequency of 600 MHz) at temperature of 25 °C. The arrows indicate the residues that presented a chemical shift perturbation upon RNA binding higher than $\Delta\delta_{ave}$ + SD. The top insert presents the titration effect in the behavior of fast exchange regime for the Asp155 upon RNA binding. The spectra were recorded at protein concentration of 100 μM (black) and at RNA concentration of 15 (blue), 30 (cyan), 55 (green), 80 (yellow), and 115 (red). (B) Chemical shift perturbation ($\Delta\delta$) for the formation of the [U-^{15}N]dgM2-1/RNA complex, $\Delta\delta = (\Delta\delta_{HN}^2 + \Delta\delta_N^2/25)$. The solid line denotes $\Delta\delta_{ave}$ and the dashed line indicates $\Delta\delta_{ave}$ + SD. The black circles are $\Delta\delta$ values of side chain of the residues Asn and Gln. The cyan stars indicate the proline residues (Pro120, Pro123, Pro153, and Pro176). The secondary structures along the sequence are indicated at the top. (C) Identification on the cdM2-1 structure of residues that present significant $\Delta\delta$ upon RNA-binding. Values of $\Delta\delta$ higher than $\Delta\delta_{ave}$ are colored in cyan: Ala73, Gly75, Val76, Val79, Ile84, Ile87, Asn88, Ile90, Thr91, Ser94, Ala95, Cys96, Ala98, Ser100, Asn142, Thr145, Leu149, Leu157, Lys158, Lys159, Ile161, and Thr164; and higher than $\Delta\delta_{ave}$ + SD are colored in green: Leu74, Lys92, Gln93, Val97, His147, Lys150, Arg151, Leu152, Asp155, and Val156. (B) Electrostatic potential surface of cdM2-1 calculated from APBS software using the charge values and protonation states (pH 6.5, 150 mM NaCl, 25 °C) determined by PDB2PQR webserver along with PROPKA program. The bar denotes the electrostatic potential range from –5 (red) to +5 kT (blue).

Figure 4. Solvent paramagnetic relaxation enhancement for free and RNA-bound cdM2-1. (A) Solvent exposure ($I_{NH}\cdot[Gd]^{-1}$) of each backbone NH amide of the protein (black square) and its complex with RNA (red) measured as a function of the residue number. [cdM2-1] = 350 μM and [RNA] = 115 μM, 25 °C, and 14.1 T (^{1}H frequency of 600 MHz). The green squares and

blue circles denote the solvent exposure for side chain of the residues Asn and Gln. The cyan stars indicate the proline residues (Pro120, Pro123, Pro153, and Pro176). The secondary structures along the sequence are indicated at the top. (B) Structural representation of hRSV cdM2-1 shown as cartoon and transparent surface. The residues Ser122, Arg126, Thr130, Ile133, and Ser137 in α4-helix (blue sticks) exhibit an $i + 4$ pattern of solvent-exposure corresponding to a nearly α-helix turn. The cyan color denotes the α1–α2–α5–α6 helix bundle.

Figure 5. Thermal susceptibility data of the amide hydrogen (^{1}HN) chemical shifts of the free and RNA-bound cdM2-1. (A) $d\delta_{HN}/dT$ values of the protein (black square) and its complex with RNA (red circle) as a function of the residue number. The temperature dependence measurements were obtained from 2D ^{1}H–^{15}H HSQC spectra at 15, 20, 25, 30, and 35 °C. [cdM2-1] = 350 μM and [RNA] = 115 μM, and 14.1 T (^{1}H frequency of 600 MHz). The green squares and blue circles denote the values of $d\delta_{HN}/dT$ for side chain of the residues Asn and Gln. The cyan stars indicate the proline residues (Pro120, Pro123, Pro153, and Pro176). The secondary structures along the sequence are indicated at the top. (B) Difference between the values of $d\delta_{HN}/dT$ for the free and RNA-bound cdM2-1 as a function of the residue number ($\Delta d\delta_{HN}/dT = d\delta_{HN}/dT_{protein} - d\delta_{HN}/dT_{protein+RNA}$). $\Delta d\delta_{HN}/dT$ + SD values far away from the zero are considered as significant changes, and the correspondent residue names are denoted. (C) The significant changes of $|d\delta_{HN}/dT|$ values between the free and bound states of the cdM2-1 are indicated in the protein structure. The values of $\Delta d\delta_{HN}/dT > 0$ (increase in $|d\delta_{HN}/dT|$) and $\Delta d\delta_{HN}/dT < 0$ (decrease in $|d\delta_{HN}/dT|$) of the ^{1}HN chemical shifts are denoted as cyan and magenta spheres, respectively. The side chain of Gln93 and Asn141 are displayed as sticks with amide hydrogens colored according the change of $\Delta d\delta_{HN}/dT$. The side chain of Asn141 forms

a hydrogen bond (black dashed line) with backbone of Ser137. Ile146 is shown as green sticks because presents $d\delta_{HN}/dT > -5.0$ ppb·K^{-1}.

Figure 6. Evidence of the structural changes of RNA induced by cdM2-1. 1D ^{1}H NMR spectra of the imino (11–15 ppm) and amine/amide (6–10 ppm) proton region of the protein-free (blue line) and bound (red line) state of RNA. The black line presents the 1D ^{1}H spectrum of the free cdM2-1 as a negative control, especially for the imino region. [cdM2-1] = 350 μM and [RNA] = 115 μM, 25 °C, and 14.1 T (^{1}H frequency of 600 MHz). In imino proton region, the thick lines denote smoothed profiles that assist the visualization of the spectra.

Figure 7. ^{15}N nuclear spin relaxation data of the free and RNA-bound cdM2-1 backbone amide by NMR on pico to nanosecond timescale. Reduced spectral density function (A) $J(0.87\omega_H)$, (B) $J(\omega_N)$, and (C) $J(0)$ of the cdM2-1 (black square) and its complex with RNA (red circle) which were calculated from the ^{15}N R_1, R_2, and hetNOE relaxation data . (D) R_2/R_1 ratio calculated from relaxation rates R_1 and R_2. The black and red lines denote the average values of $J(0.87\omega_H)$, $J(\omega_N)$, and $J(0)$ for the free and RNA-bound cdM2-1, respectively. To determine the average values, those corresponding to the N and C-terminal residues were excluded. The cyan stars indicate the proline residues (Pro120, Pro123, Pro153, and Pro176). The secondary structures along the sequence are indicated at the top. Plots of the correlations of $J(0.87\omega_H)$ & $J(0)$ and $J(\omega_N)$ & $J(0)$ for the free (E and G) and RNA-bound cdM2-1 (F and H). [cdM2-1] = 350 μM, [RNA] = 115 μM, 25 °C, and 14.1 T (^{1}H frequency of 600 MHz). The red line denotes the linear fitting adjustment and R indicates the correlation coefficient.

Figure 8. The extended model-free formalism of Lipari-Szabo applied to analyze the nuclear spin relaxation data. (A) Order parameter (S^2) and (B) conformational exchange rate

(R_{ex}) calculated by TENSOR2 software from ^{15}N nuclear spin relaxation data of the backbone of cdM2-1 using the Lipari–Szabo model free formalism. The continuous and dashed red line in (A) denote the average value of S^2 (0.91 ± 0.04), excluding terminal residues, and S^2 = 0.7, respectively. The dashed red line in (B) depicts the cutoff value of R_{ex} equal to 1.0 Hz. [cdM2-1] = 350 μM, [RNA] = 115 μM, 25 °C, and 14.1 T (^1H frequency of 600 MHz). Cartoon representation of cdM2-1 highlighted in blue and red the residues that showed significant values of (C) S^2 (< 0.91 ± 0.04; Ile90, Asp110, Lys113, Leu114, Glu119, Asn121, Ser122, and Asn141) and (D) R_{ex} (> 1.0 Hz; Glu59, Ile84, Ile87, Asn88, Ile90, Glu119, and Ala154), respectively.

Figure 9. Analysis of the molecular dockings of the structural models of the cdM2-1/RNA complex. (A) Structural models of the cdM2-1/RNA complex calculated by 3dRPC webserver. The backbone of the protein (colorful) and 10 nucleic acids (orange) are presented as cartoon and ribbon, respectively. The protein also is shown as surface representation and its helices are highlighted using different colors (α1: red, α2: green, α3: yellow, α4: blue, α5: purple, α6: cyan). (B) Structural model of cdM2-1/RNA complex which exemplified the preferential binding of the protein (as grey cartoon) close to secondary structural regions of the RNA molecules (as cartoon: backbone in orange and nucleoside in green/blue). (C) Mass density map (green mesh with isovalue of 0.5) of 10 RNA molecules bound to the protein (as grey cartoon) generated by VMD programs using the VolMap tool. (D) Absolute accessible surface area change (ΔASA) of the cdM2-1 (as colorful worm) bound to the nucleic acids. The thickness and colors represent the degree of ΔASA, with the strongly and weakly accessible residues colored in blue (thin) and red (thick), respectively. The structural representations were made in PyMol and VMD programs.

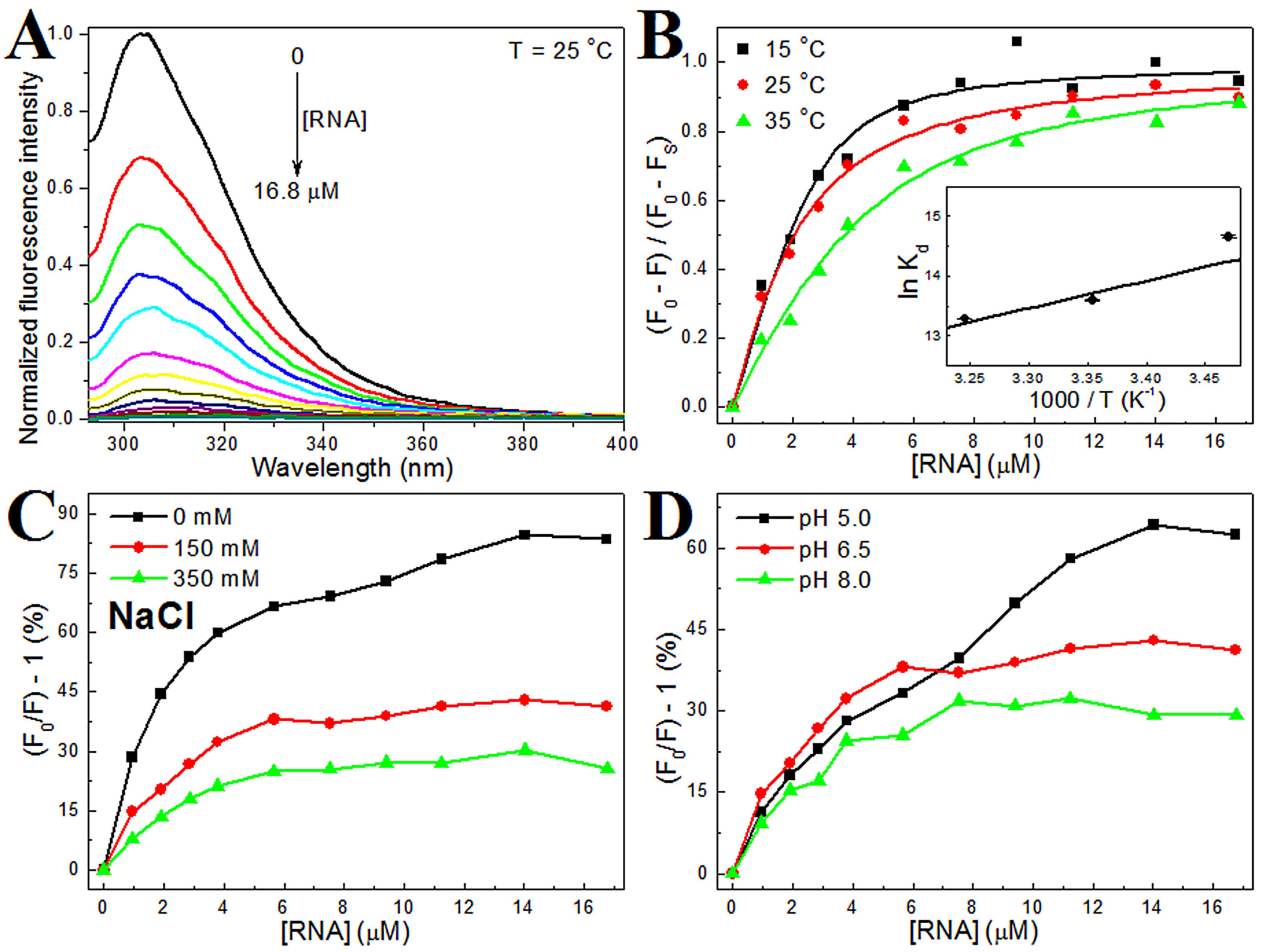
Figure 10. Analysis of the molecular dynamics of the structural models of the cdM2-1/RNA complex. (A) RMSD values for the non-hydrogen atoms of RNAs (green line), for the backbone atoms of the helical-core region of the protein (cdM2-$1_{75–171}$, black line), and for the backbone

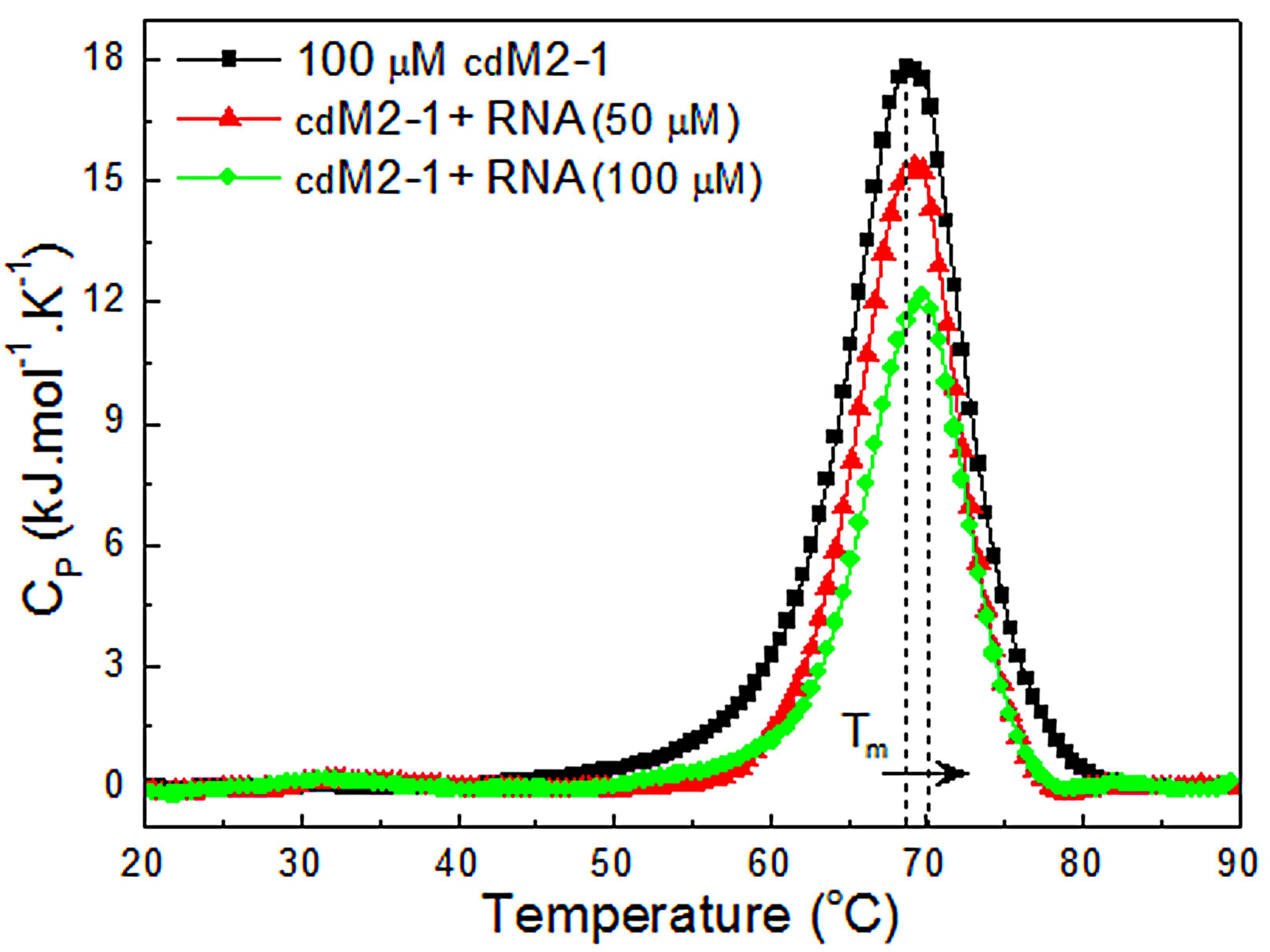
atoms of the entire cdM2-1 (red line). (B) Number of contacts between atoms of cdM2-$1_{75–171}$ and RNAs for distances < 0.6 nm. (C) Number of hydrogen bonds formed between the RNA molecules and the cdM2-$1_{75–171}$ (black line), the residue Lys150 (red line), and residue Arg151 (green line).

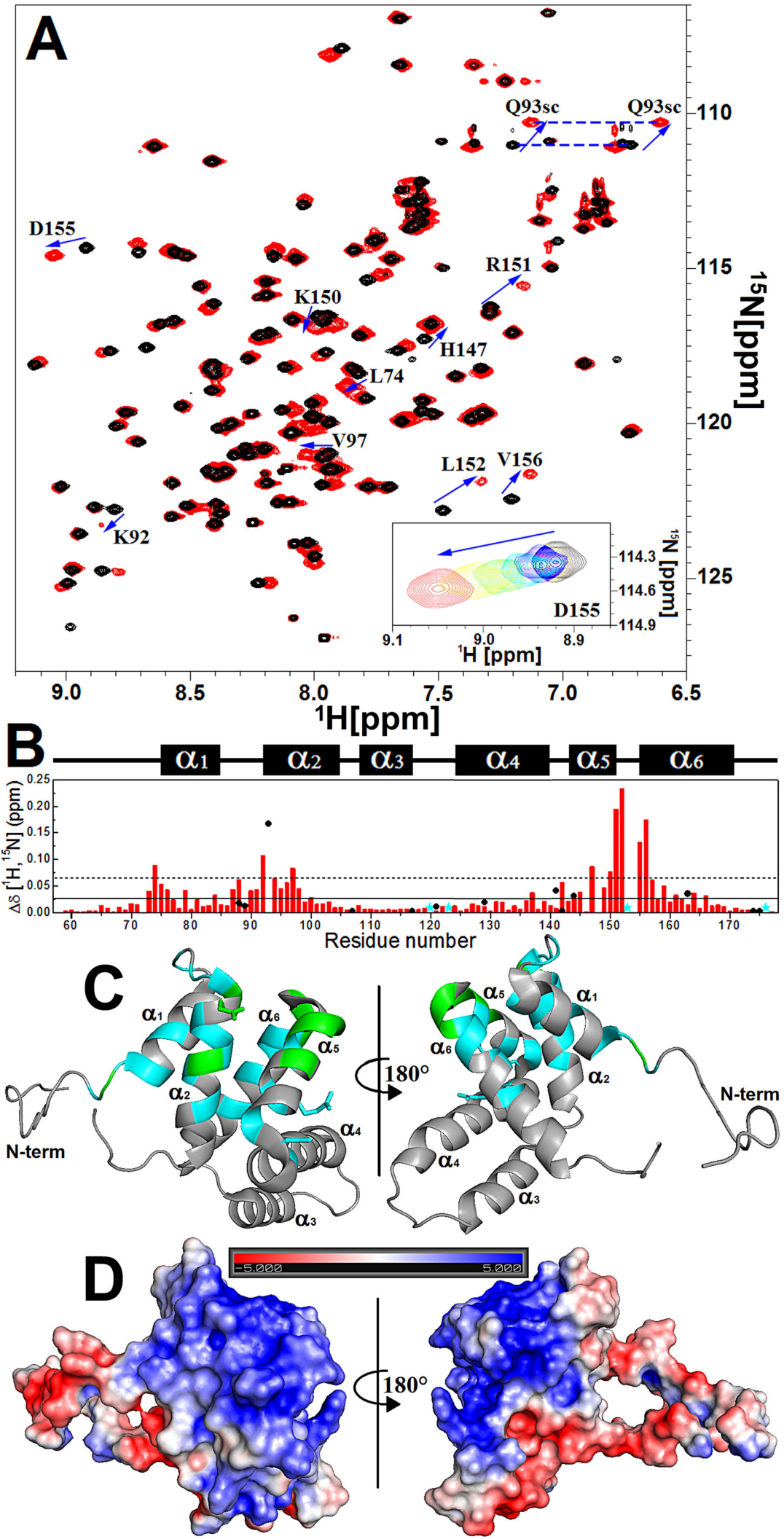
Table 1. Dissociation constant (K_d), stoichiometry coefficient (n_{SC}), enthalpy change (ΔH), Gibbs free energy change (ΔG), and entropy change (ΔS) of the cdM2-1/RNA interaction determined using fluorescence quenching experiments at 15, 25, and 35 °C.[a]

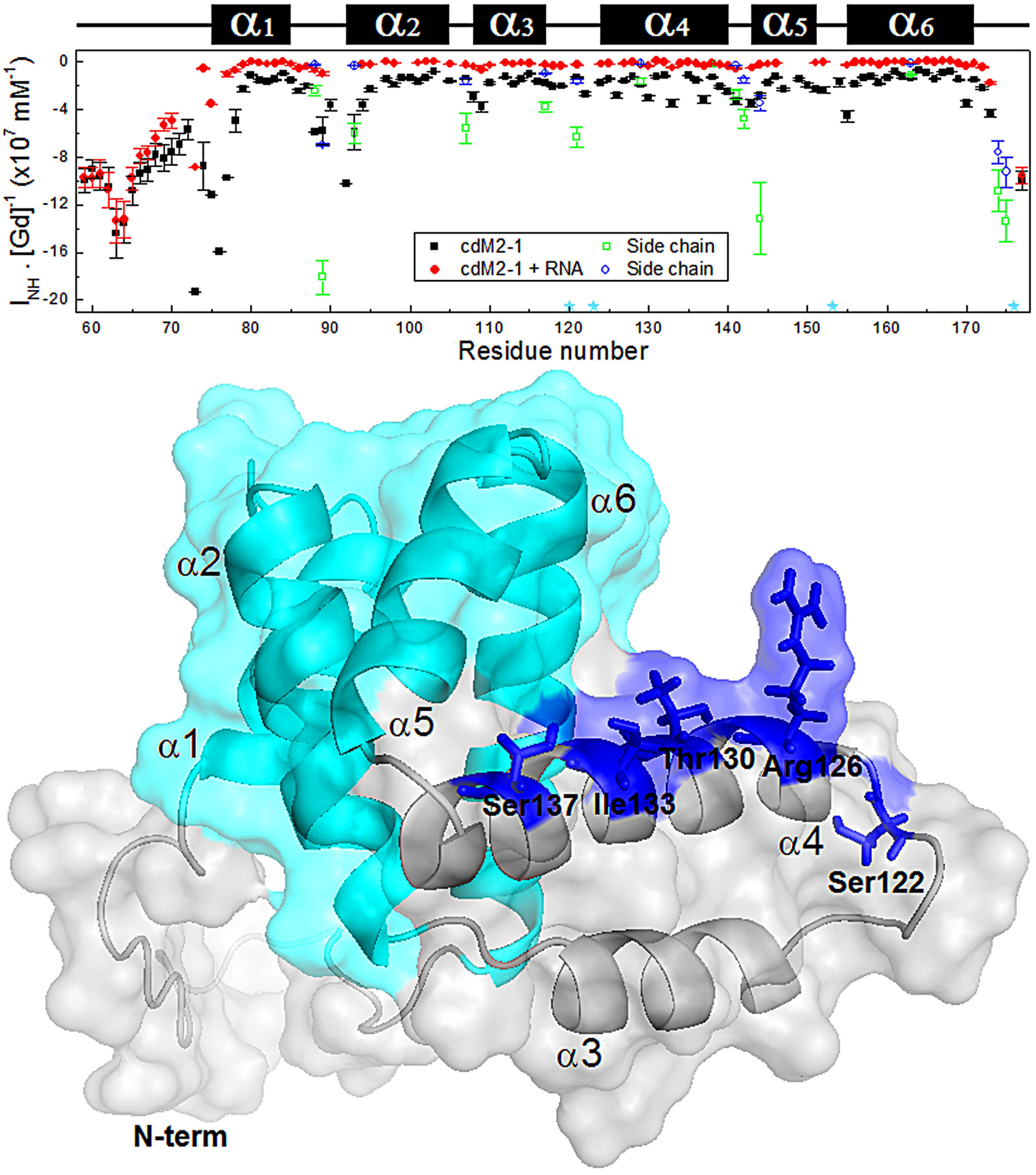
$T(°C)$	K_d ($\times 10^{-6}$ M)	n_{SC}	$\Delta H°$ ($kJmol^{-1}$)	$\Delta G°$ ($kJmol^{-1}$)	$\Delta S°$ ($Jmol^{-1}K^{-1}$)
15	0.43 ± 0.18	2.0 ± 0.4		–35.1 ± 0.7	–34 ± 35
25	1.24 ± 0.18	3.3 ± 1.1	–45 ± 10	–33.7 ± 0.4	–37 ± 34
35	1.70 ± 0.28	1.4 ± 0.2		–34.0 ± 0.4	–36 ± 32

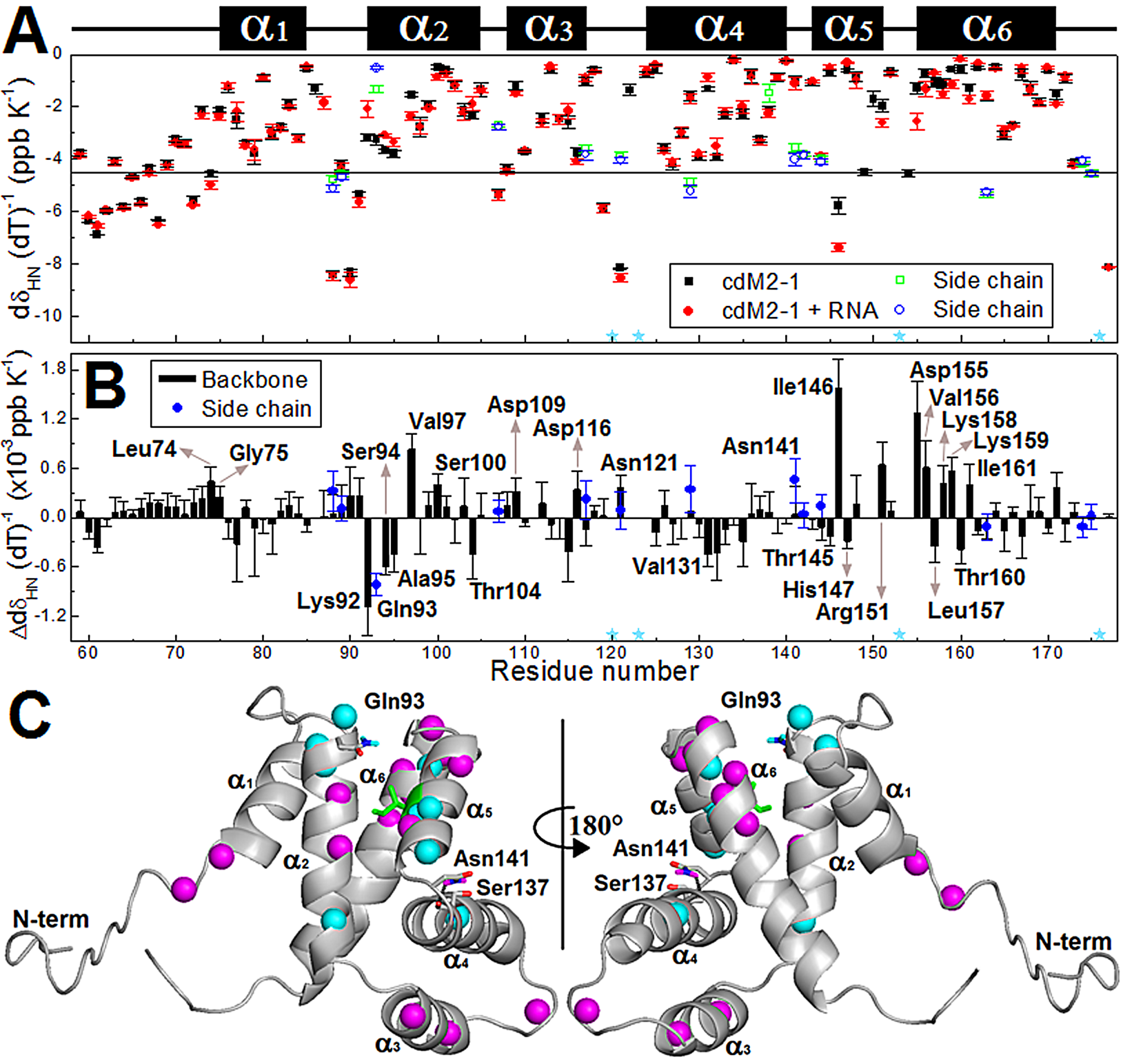
[a]All experiments were performed in 50 mM phosphate buffer (pH 6.5) containing 150 mM NaCl and 1.0 mM DTT.

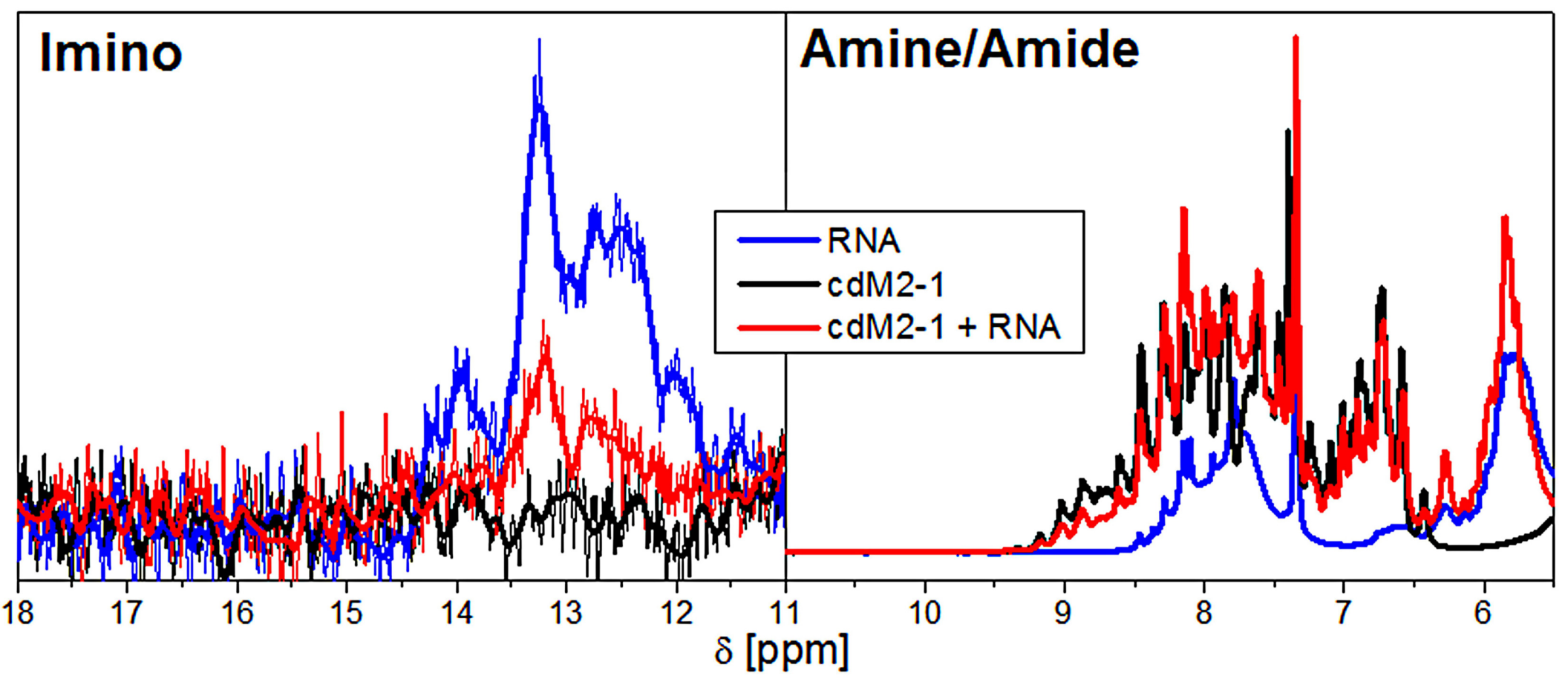

A
T = 25 °C
0
[RNA]
16.8 μM
Normalized fluorescence intensity
Wavelength (nm)
B
15 °C
25 °C
35 °C
(F0 - F) / (F0 - Fs)
[RNA] (μM)
ln Kd
1000 / T (K-1)
C
0 mM
150 mM
350 mM
NaCl
(F0/F) - 1 (%)
[RNA] (μM)
D
pH 5.0
pH 6.5
pH 8.0
(F0/F) - 1 (%)
[RNA] (μM)

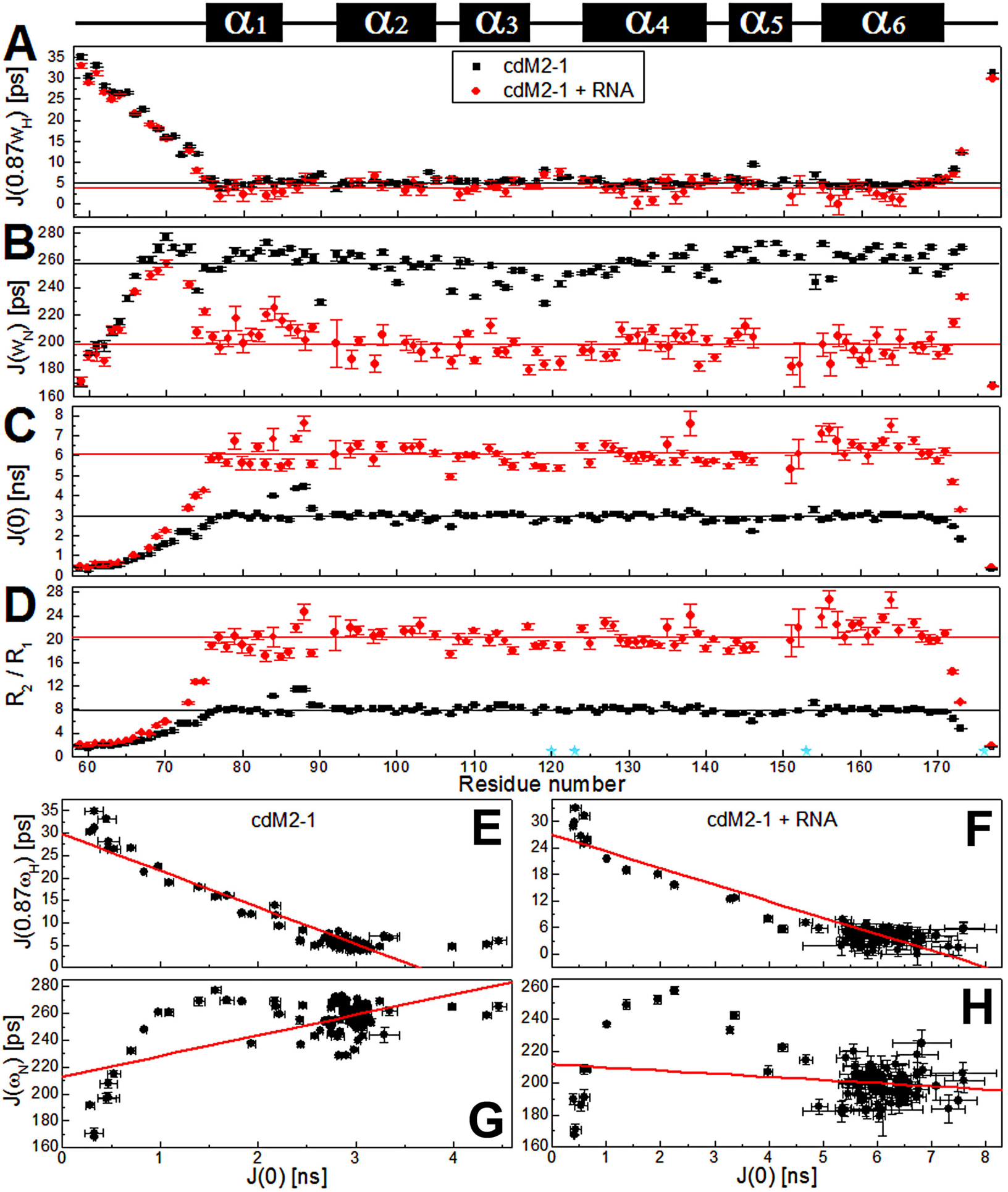

100 μM cdM2-1
cdM2-1+ RNA(50 μM)
cdM2-1+ RNA(100 μM)
C_P (kJ.mol^{-1}.K^{-1})
T_m
Temperature (°C)
0
3
6
9
12
15
18
20
30
40
50
60
70
80
90

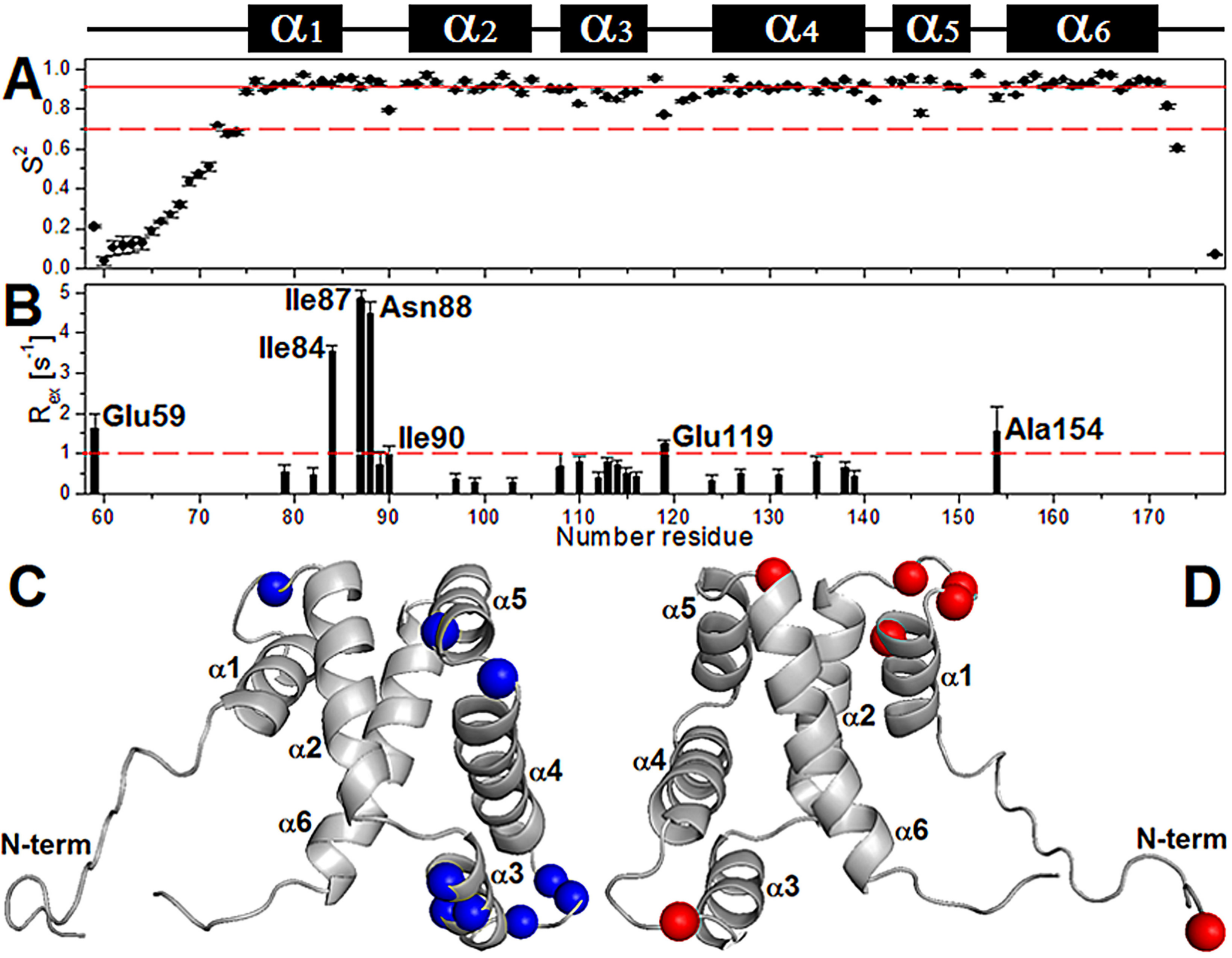

A
Q93sc
Q93sc
D155
R151
K150
H147
L74
V97
L152
V156
K92
D155
114.3
114.6
114.9
9.1
9.0
8.9
^{1}H [ppm]
^{15}N [ppm]
110
115
120
125
9.0
8.5
8.0
7.5
7.0
6.5
^{1}H[ppm]
^{15}N[ppm]
B
α1
α2
α3
α4
α5
α6
Δδ [^{1}H, ^{15}N] (ppm)
0.25
0.20
0.15
0.10
0.05
0.00
60
70
80
90
100
110
120
130
140
150
160
170
Residue number
C
α1
α6
α5
α2
α4
α3
N-term
180°
α5
α1
α6
α2
N-term
α4
α3
D
-5.000
5.000
180°

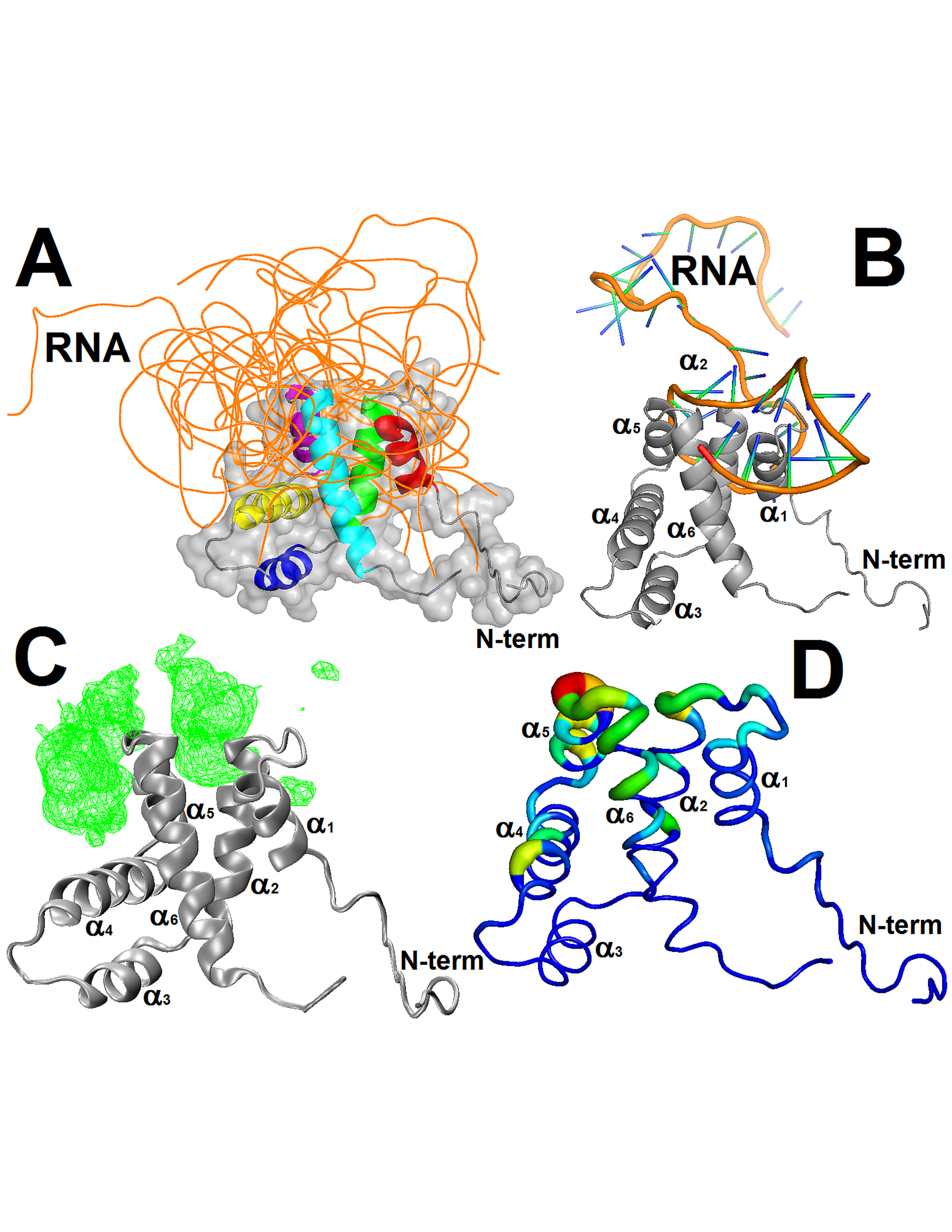

α1
α2
α3
α4
α5
α6
$I_{NH} \cdot [Gd]^{-1}$ ($x10^7$ mM^{-1})
0
-4
-8
-12
-16
-20
60
70
80
90
100
110
120
130
140
150
160
170
Residue number
cdM2-1
cdM2-1 + RNA
Side chain
Side chain
α6
α2
α5
α1
Thr130
Arg126
Ser137
Ile133
α4
Ser122
α3
N-term

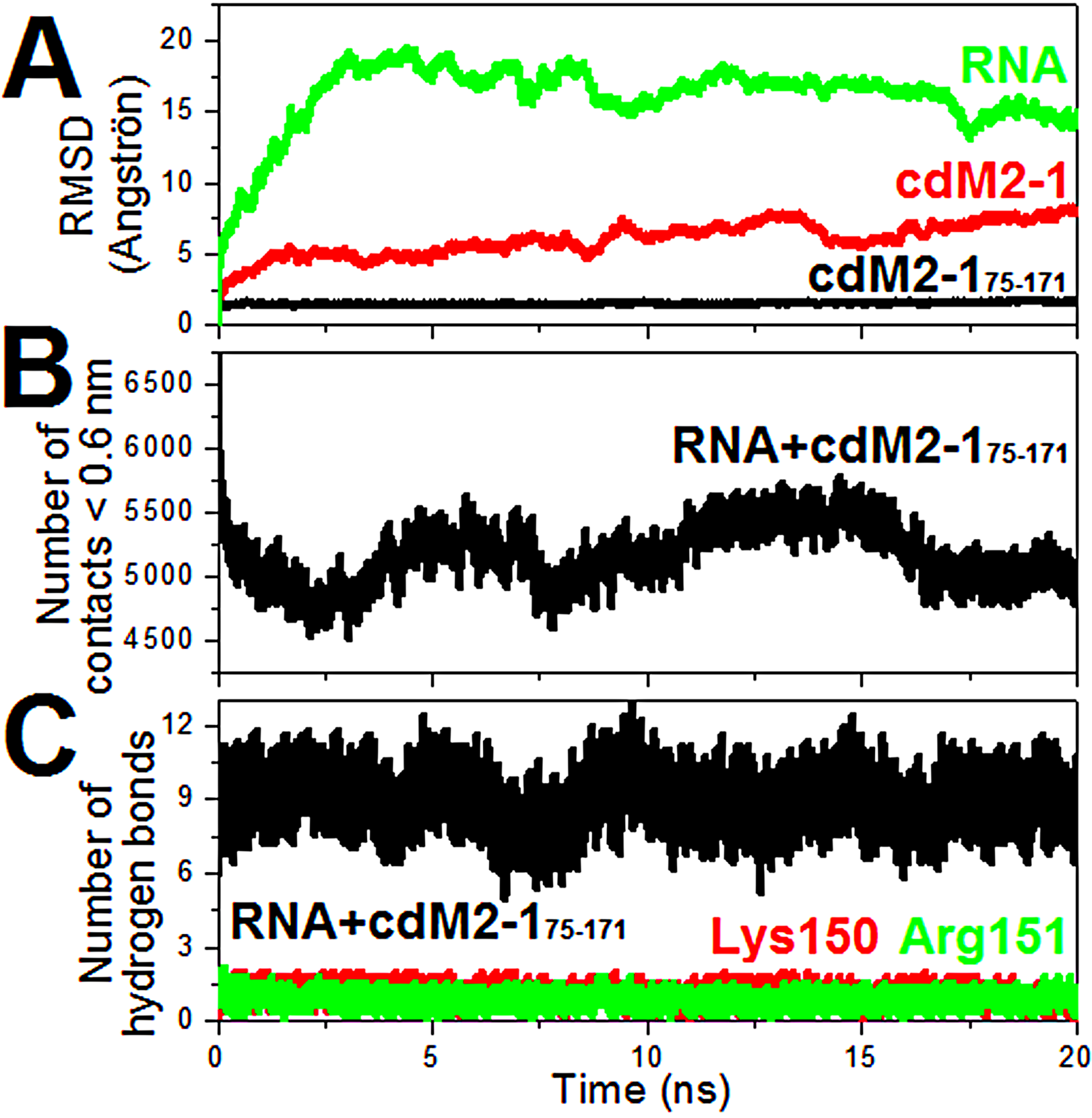

A
α1
α2
α3
α4
α5
α6
dδHN (dT)-1 (ppb K-1)
cdM2-1
cdM2-1 + RNA
Side chain
Side chain
B
Backbone
Side chain
ΔdδHN (dT)-1 (x10-3 ppb K-1)
Leu74
Gly75
Ser94
Val97
Ser100
Asp109
Asp116
Asn121
Asn141
Ile146
Asp155
Val156
Lys158
Lys159
Ile161
Lys92
Gln93
Ala95
Thr104
Val131
Thr145
His147
Arg151
Leu157
Thr160
Residue number
C
Gln93
α1
α6
α5
α2
Asn141
Ser137
α4
α3
N-term
180°
Gln93
α6
α5
α1
Asn141
α2
Ser137
α4
α3
N-term


Imino
Amine/Amide
RNA
cdM2-1
cdM2-1 + RNA
18
17
16
15
14
13
12
11
10
9
8
7
6
δ [ppm]


α1
α2
α3
α4
α5
α6
A
B
C
D
cdM2-1
cdM2-1 + RNA
J(0.87ω_H) [ps]
J(ω_N) [ps]
J(0) [ns]
R_2 / R_1
Residue number
E
F
G
H
cdM2-1
cdM2-1 + RNA
J(0.87ω_H) [ps]
J(ω_N) [ps]
J(0) [ns]
J(0) [ns]


α1
α2
α3
α4
α5
α6
A
S2
1.0
0.8
0.6
0.4
0.2
0.0
B
Rex [s-1]
Ile87
Asn88
Ile84
Glu59
Ile90
Glu119
Ala154
0
1
2
3
4
5
60
70
80
90
100
110
120
130
140
150
160
170
Number residue
C
α5
α1
α2
α4
α6
α3
N-term
D
α5
α1
α2
α4
α6
α3
N-term


A
RNA
N-term
B
RNA
α2
α5
α4
α6
α1
N-term
α3
C
α5
α1
α2
α4
α6
N-term
α3
D
α5
α1
α2
α6
α4
N-term
α3


A
RMSD (Angström)
20
15
10
5
0
RNA
cdM2-1
cdM2-1 75-171
B
Number of contacts < 0.6 nm
6500
6000
5500
5000
4500
RNA+cdM2-1 75-171
C
Number of hydrogen bonds
12
9
6
3
0
RNA+cdM2-1 75-171
Lys150
Arg151
0
5
10
15
20
Time (ns)