Supplementary Information

The human RNA-binding protein and E3 ligase MEX-3C binds the MEX-3–recognition element (MRE) motif with high affinity

Lingna Yang[1], Chongyuan Wang[1], Fudong Li[1], Jiahai Zhang[1], Anam Nayab[1], Jihui Wu[1], Yunyu Shi[1, 2] and Qingguo Gong[1,*]

[1]Hefei National Laboratory for Physical Science at Microscale, Collaborative Innovation Center of Chemistry for Life Sciences and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230027, China
[2] CAS Center for Excellence in Biomacromolecules, **C**hinese Academy of Sciences, Beijing, 100101, China

Running title: *Molecular basis of RNA recognition by MEX-3C*

*To whom correspondence should be addressed. Tel: +86 551 63607644; Fax: +86 551 63601443; Email: qgg@ustc.edu.cn

Keywords: MEX-3C, KH domain, RNA, crystal structure, E3 ubiquitin ligase, RNA-protein complex, RNA binding protein, RNA-protein interaction

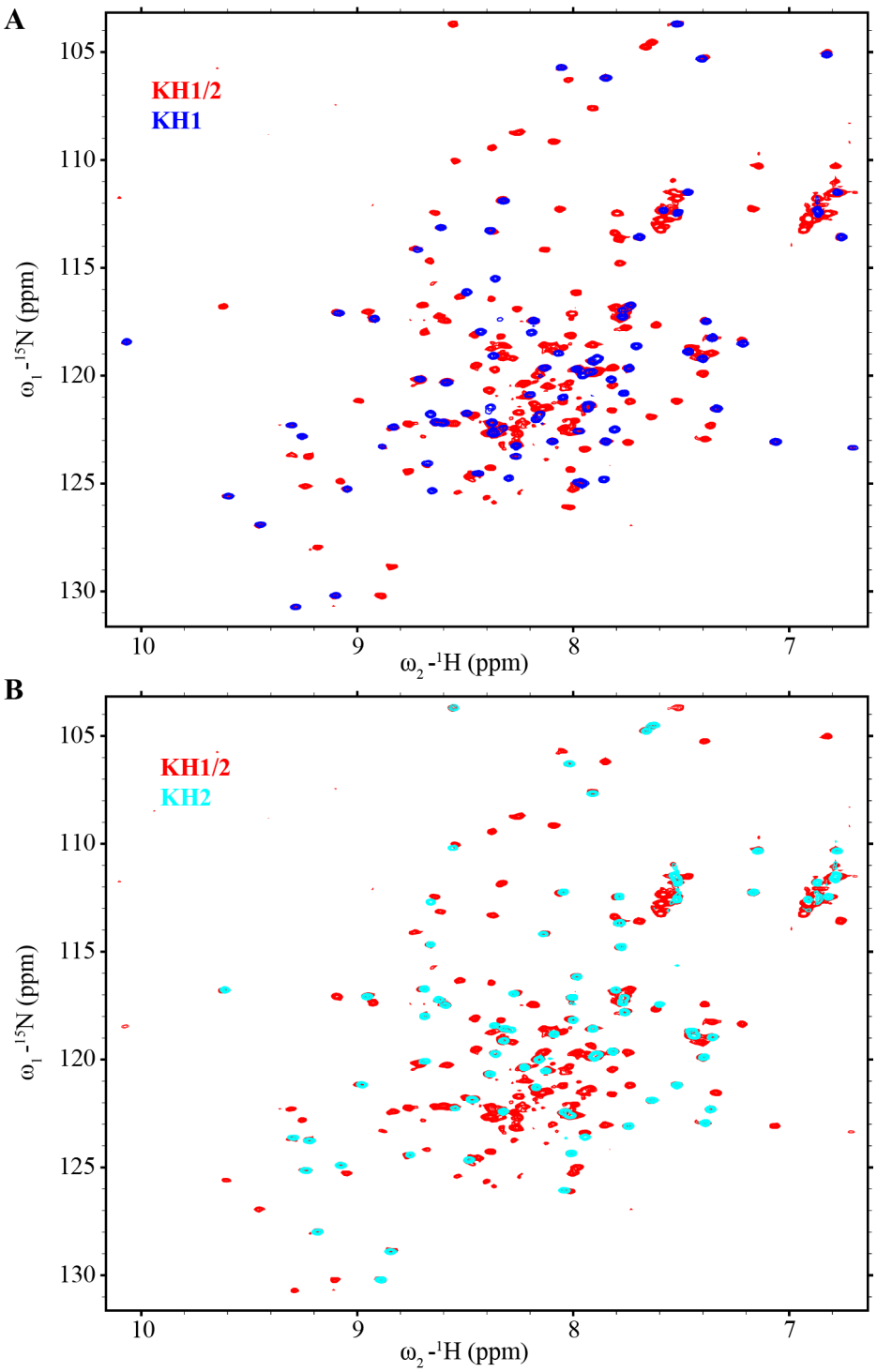

Figure S1. Superposition of the ^{1}H-^{15}N HSQC spectra of KH1–KH1/2 (A) and KH2–KH1/2 (B). The NMR peaks of the individual KH1 and KH2 domains can be aligned well with those in the KH1/2 domains. Note that the KH1/2 domains has a linker region of 16 residues between the KH1 and KH2 domains.

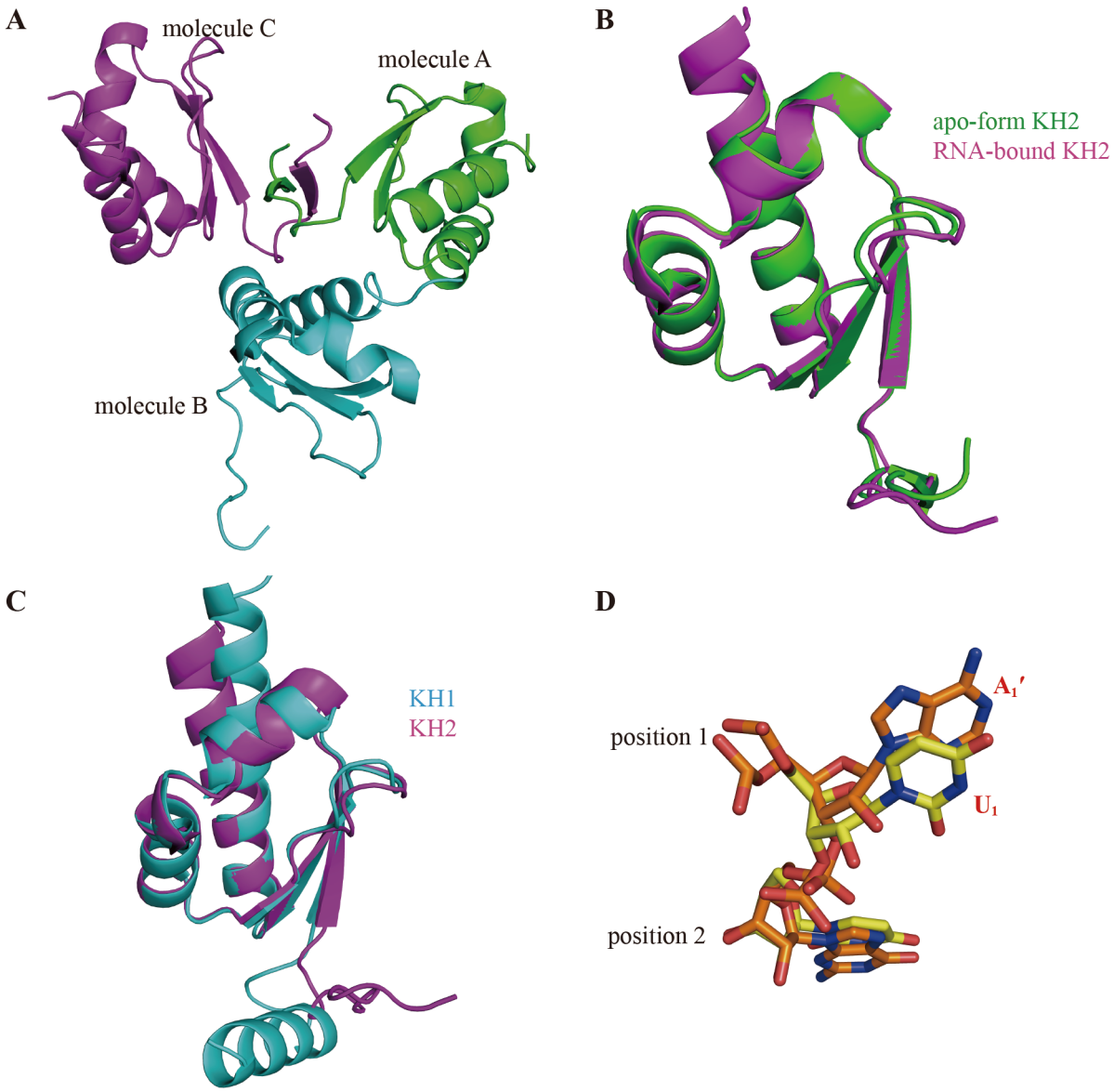

Figure S2. Structures of the hMEX-3C KH domains. (A) Structure of the apo-form KH2 is shown in cartoon. Three copies of the KH2 domain were identified in one asymmetric unit. The molecule A and C formed as a dimer due to crystal packing. (B) Structural alignment of the apo and RNA-bound forms of KH2 domain. (C) Structural alignment of the hMEX-3C KH1 and KH2 domains. (D) U_1 base ring in the KH1 complex swings away from the position of the A_1' base in the KH2 complex towards position 2.

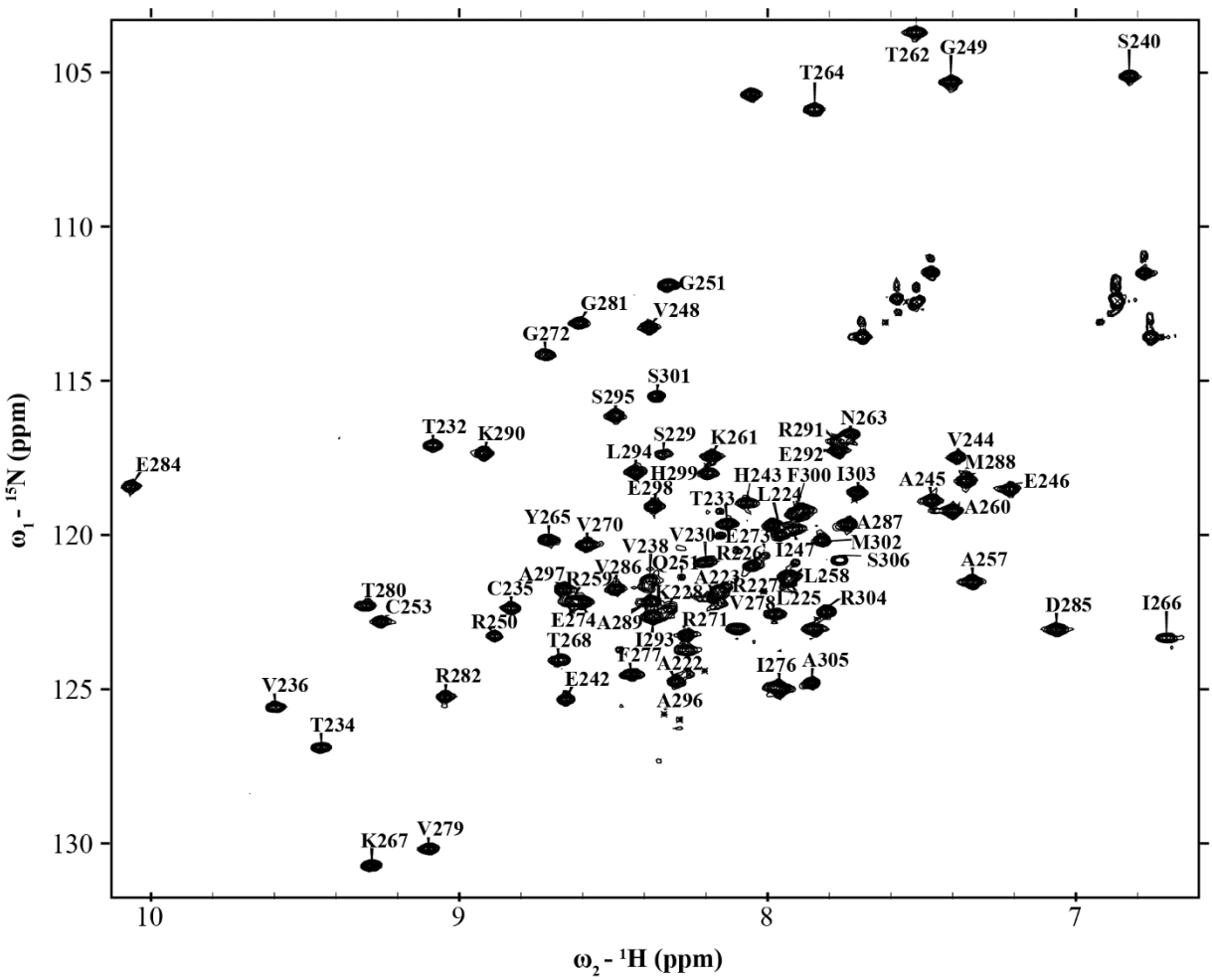

Figure S3. Backbone assignment of the hMEX-3C KH1 domain in the ^{1}H-^{15}N HSQC spectrum.

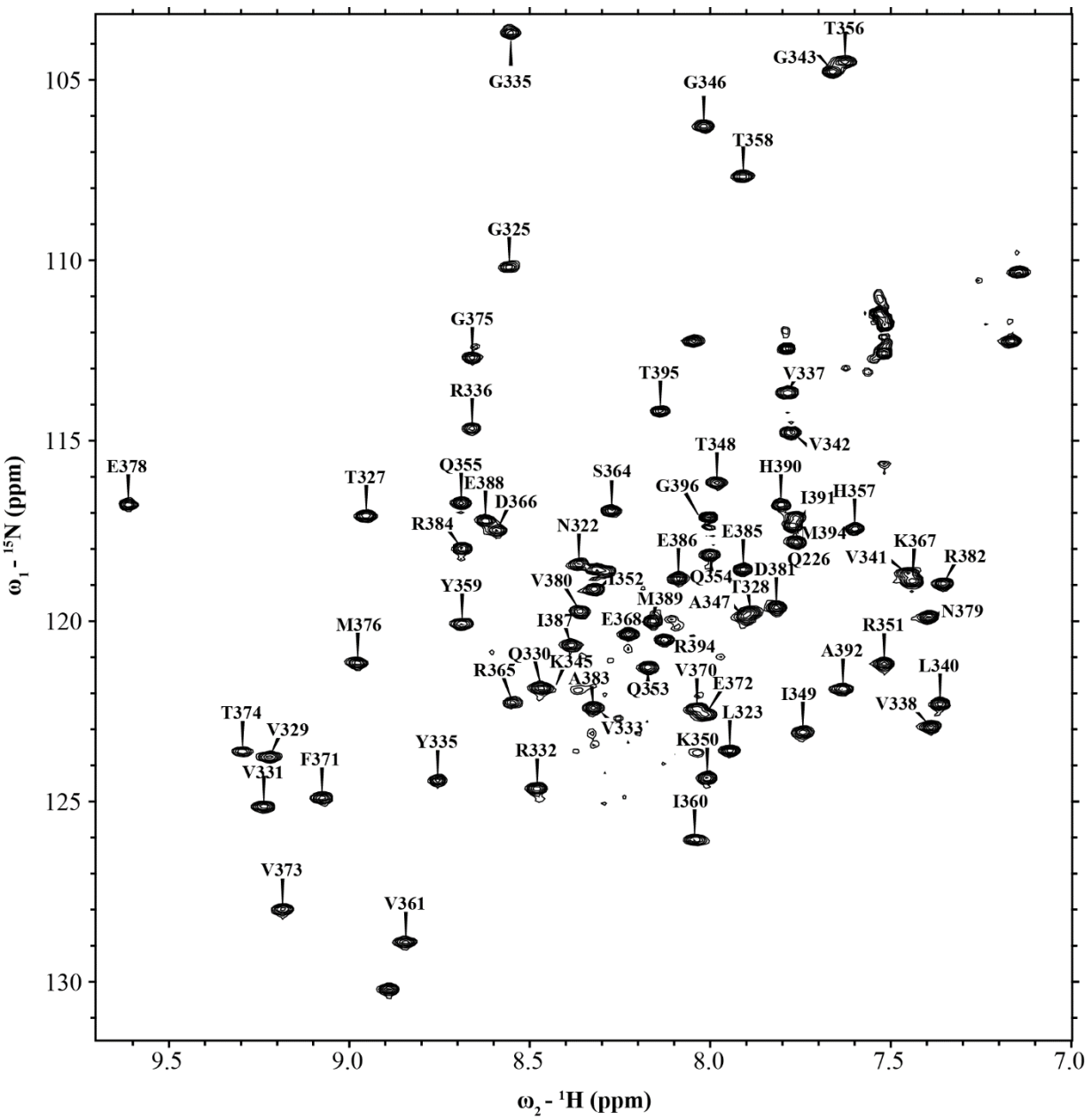

Figure S4. Backbone assignment of the hMEX-3C KH2 domain in the ^{1}H-^{15}N HSQC spectrum.

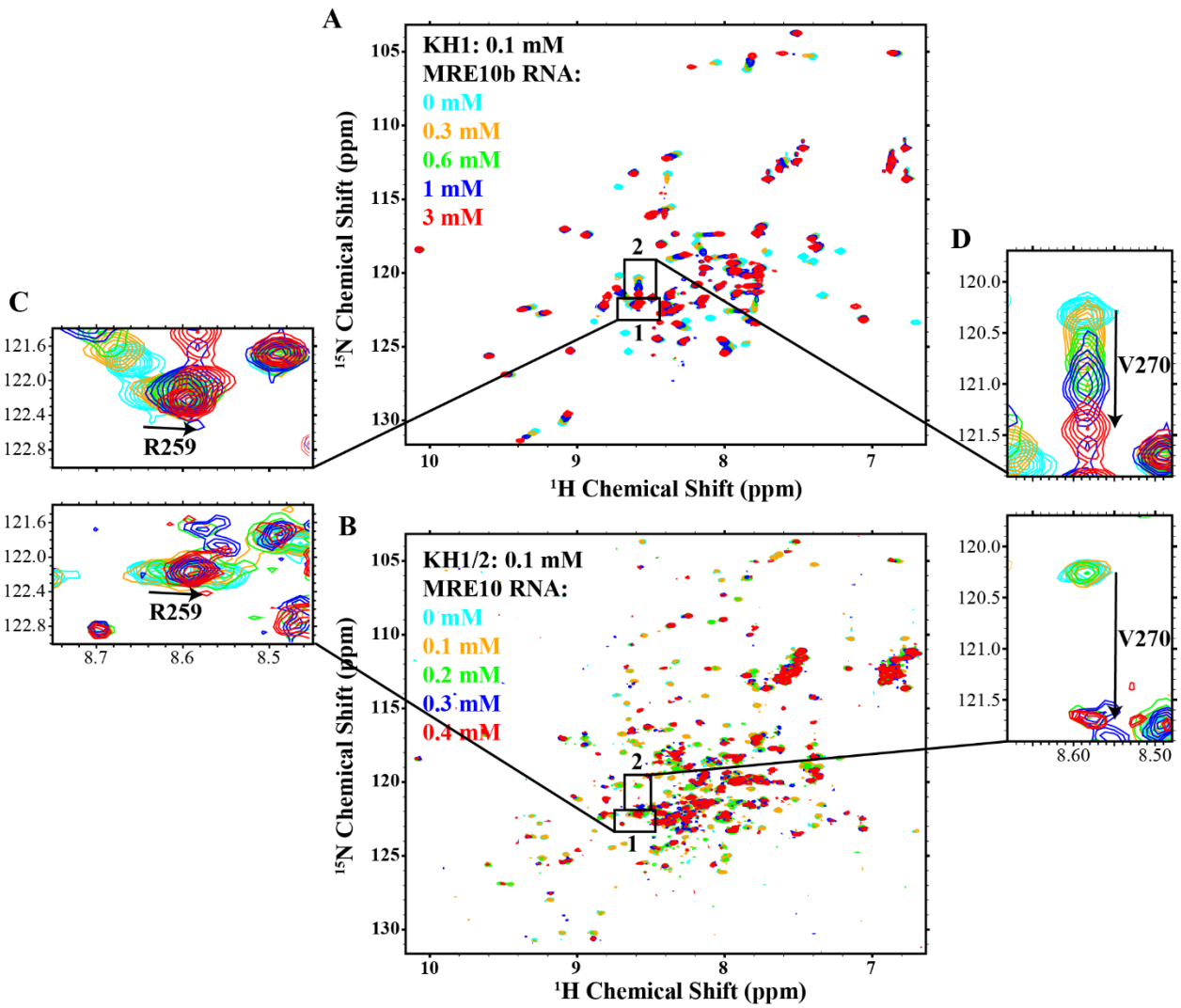

Figure S5. ^{1}H-^{15}N HSQC chemical shift perturbation of the KH1 domain by MRE10b (GUUUAG) (A) and the KH1/2 domains by MRE10 (CAGAGUUUAG) (B). (C, D) Several residues located at the KH1-RNA binding interface show similar modes of chemical shift perturbation in both the individual KH1 domain and the KH1/2 domains.

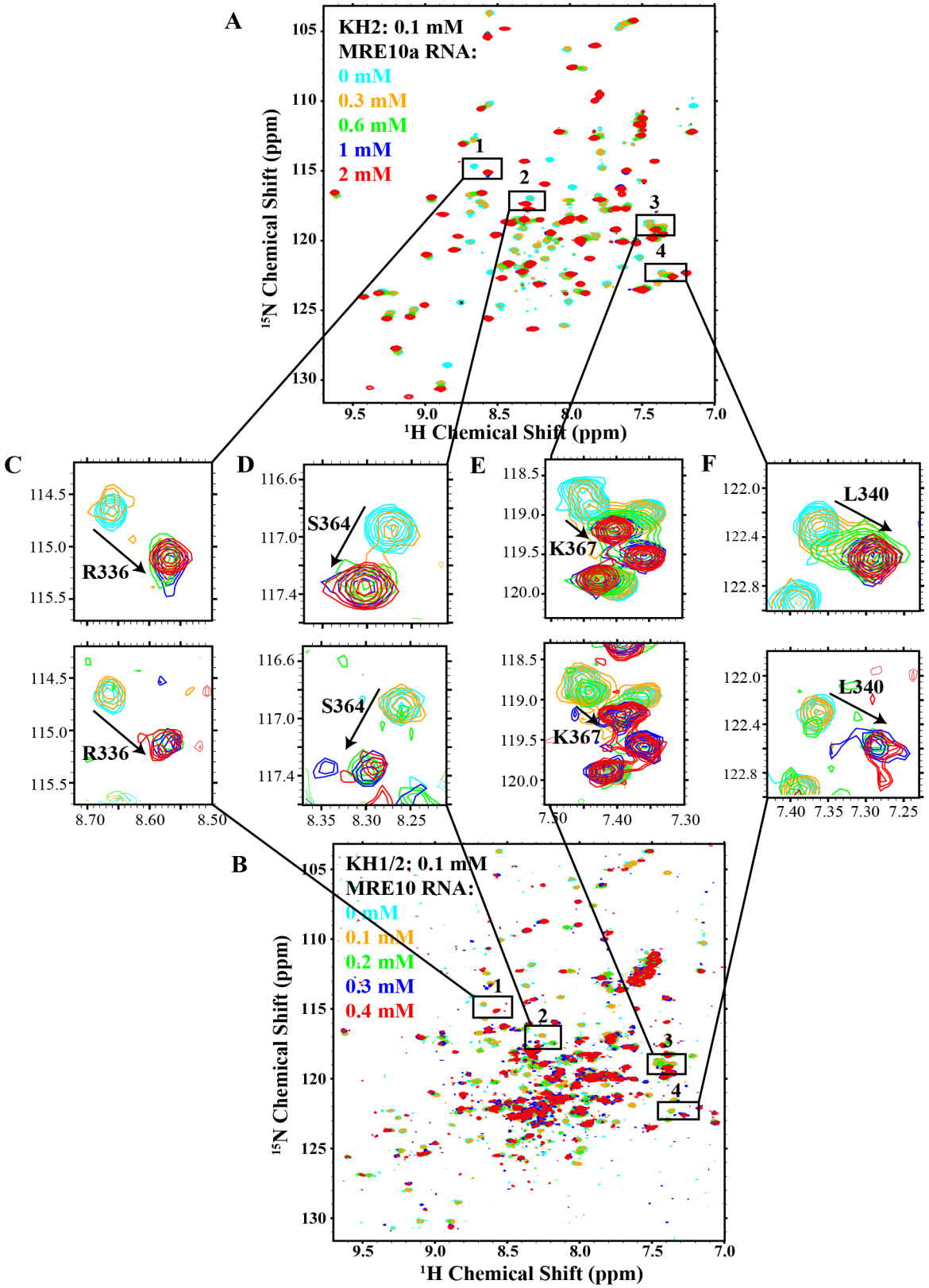

Figure S6. ^{1}H-^{15}N HSQC chemical shift perturbation of the KH2 domain by MRE10a (CAGAGU) (A) and the KH1/2 domains by MRE10 (CAGAGUUUAG) (B). (C-F) Several residues located at the KH2-RNA binding interface show similar modes of chemical shift perturbation in both the individual KH2 domain and the KH1/2 domains.

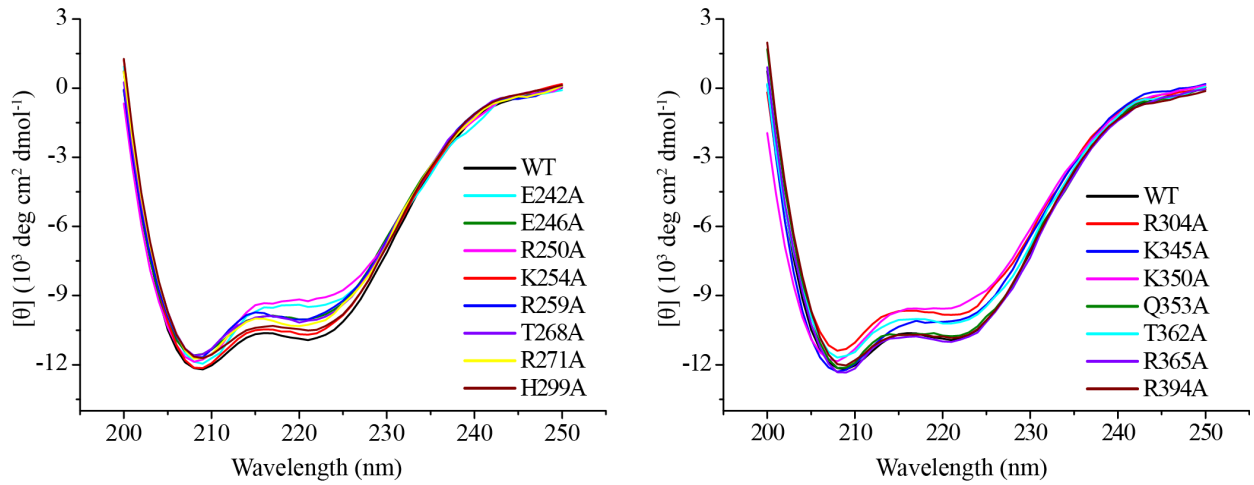

Figure S7. CD spectra of hMEX-3C KH1/2 domains and its mutants.

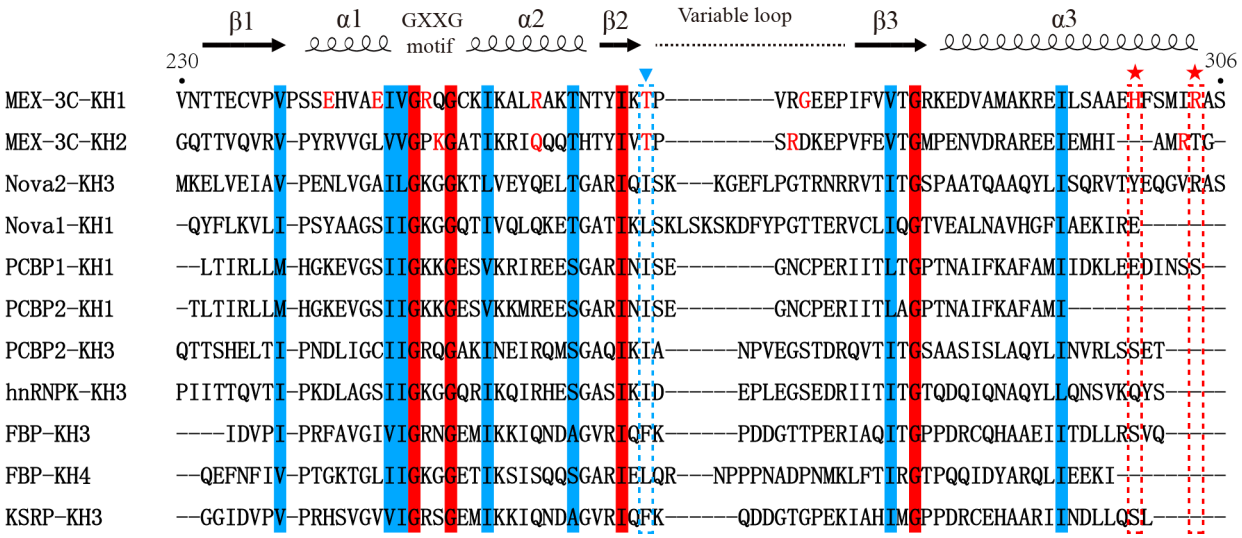

Figure S8. Sequence alignment of human MEX-3C KH domains and other KH domains with reported complex structures. Conserved and similar residues are highlighted in red and blue, respectively. The residues involved in hydrogen-bonding interactions via their side chains are colored in red. The T268/T362 residues involved in the formation of H bonds with the Hoogsteen edge of the adenine at position 3 are indicated by blue inverted triangle. The H299 and R304 residues in the α3 helix of KH1 are indicated by red asterisks. The secondary structure elements of the KH domain assigned on the basis of the structure of hMEX-3C KH1 domain are shown at the top.

Table S1. Data collection and refinement statistics.

	MEX-3C KH1-GUUUAG	MEX-3C KH2-CAGAGU	MEX-3C KH2
PDB ID	5WWW	5WWX	5WWZ
Data Collection			
Space group	$P6_2$	$P4_32_12$	$P4_22_12$
Cell dimensions			
a, b, c (Å)	76.37, 76.37, 36.39	74.61, 74.61, 32.53	83.32, 83.32, 78.96
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 90
Wavelength (Å)	0.9776	0.9776	0.9776
Resolution* (Å)	38.18-1.80 (1.86-1.80)	33.37-2.0 (2.07-2.0)	37.26-2.50 (2.59-2.50)
Completeness (%)	99.5(98.9)	99.6(99.8)	100.0(100.0)
Redundancy	10.1(10.3)	7.6(7.9)	6.9(6.9)
R_{merge} (%)	15.9(70.5)	10.8(53.4)	8.2(56.3)
$I/\sigma I$	19.1(2.8)	11.8(2.2)	15.7(2.3)
Refinement			
No. reflections used/free	11402/564	6586/337	10107/467
Resolution (Å)	38.18-1.80	33.37-2.0	37.26-2.50
R_{work}/R_{free}	17.2/20.9	17.0/21.8	22.6 /27.0
R.m.s.deviations			
Bondslengths (Å)	0.006	0.008	0.002
Bond angles (°)	0.78	0.92	0.43
B-factors (Å^2)			
Protein	24.32	22.36	56.77
RNA	30.78	22.03	N/A
Water	36.13	32.70	46.64
No. atoms			
Protein	759	703	1795
RNA	125	110	N/A
Water	93	62	40
Ramachandran plot			
Favored/allowed/outlier (%)	99.0/1.0/0	98.8/1.2/0	98.27/1.73/0

Statistics for the highest-resolution shell are shown in parentheses.

Table S2. The ITC thermodynamic parameters of the ITC experiments.

Proteins	RNA 5'-3'	ΔH kcal/mol	ΔS cal/mol/K	K_d μM	N
MEX-3C KH1/2	CAGAGUUUAG	-22.36	-45.3	0.17 ± 0.01	1.21
	C***G***GAGUUUAG	-19.1	-35.7	0.36 ± 0.02	1.19
	C***C***GAGUUUAG	-16.3	-26.9	0.52 ± 0.05	1.25
	C***U***GAGUUUAG	-15.5	-23.4	0.35 ± 0.01	1.4
	CA***A***AGUUUAG	-15.3	-23.6	0.58 ± 0.02	1.38
	CA***C***AGUUUAG	-21.3	-43.9	0.56± 0.02	1.12
	CA***U***AGUUUAG	-18.6	-34.1	0.40 ± 0.04	1.14
	CAG***G***GUUUAG	-16.0	-27.7	1.48 ± 0.16	1.18
	CAG***C***GUUUAG			ND	
	CAG***U***GUUUAG	-18.2	-35.5	1.69 ± 0.10	1.18
	CAGA***A***UUUAG	-19.4	-36.9	0.39 ± 0.04	1.14
	CAGA***C***UUUAG	-19.8	-39.8	0.89 ± 0.09	1.26
	CAGA***U***UUUAG	-19.9	-38.3	0.37 ± 0.03	1.14
	CAGAG***A***UUAG	-17.8	-30.1	0.21 ± 0.02	1.24
	CAGAG***G***UUAG	-18.9	-33.9	0.22 ± 0.01	1.34
	CAGAG***C***UUAG			ND	
	CAGAGU***A***UAG	-18.9	-33.4	0.15 ± 0.01	1.27
	CAGAGU***G***UAG	-19.4	-37.3	0.46 ± 0.05	1.36
	CAGAGU***C***UAG			>10	
	CAGAGUU***A***AG	-19.2	-35.2	0.23 ± 0.02	1.31
	CAGAGUU***G***AG	-13.7	-19.6	1.08 ± 0.08	1.25
	CAGAGUU***C***AG	-11.2	-10.8	1.18 ± 0.12	1.36
	CAGAGUUU***G***G	-14.3	-18.9	0.38 ± 0.03	1.01
	CAGAGUUU***C***G	-15.3	-22.5	0.44 ± 0.05	0.86
	CAGAGUUU***U***G	-17.5	-30.0	0.35 ± 0.03	1.37
	CAGAGUUUA***A***	-16.5	-25.1	0.17 ± 0.01	1.41
	CAGAGUUUA***C***	-21.6	-43	0.20 ± 0.01	1.24
	CAGAGUUUA***U***	-17.9	-30.9	0.24 ± 0.01	1.37
	CAGAG***C3***UUUAG	-12.83	-13.4	0.24 ± 0.02	1.25
	CAGAG***C6***UUUAG	-16.98	-28.3	0.34 ± 0.04	1.18
	CAGAG***C7***UUUAG	-15.3	-22.9	0.38 ± 0.04	0.80
	CAGAG***C8***UUUAG	-14.14	-18.7	0.35 ± 0.02	1.39
	CAGAG***C9***UUUAG	-15.85	-25.9	0.70 ± 0.07	1.22
	CAGAG***C12***UUUAG	-17.53	-32	0.83 ± 0.07	1.17
E242A (KH1/2)	CAGAGUUUAG	-20.27	-39.7	0.36 ± 0.01	1.33
E246A		-21.96	-45.0	0.30 ± 0.02	1.22
R250A		-24.45	-52.5	0.21 ± 0.01	1.10
K254A		-18.21	-32.7	0.38 ± 0.02	1.40
R259A		-23.05	-47.9	0.20 ± 0.01	1.30
T268A		-16.63	-29.2	1.03 ± 0.07	1.27
R271A		-20.40	-41.4	0.71 ± 0.05	1.14
H299A		-17.96	-32.5	0.52 ± 0.02	1.35
R304A		-21.40	-44.5	0.59 ± 0.02	1.23
K345A		-26.72	-60.2	0.37 ± 0.03	0.96

K350A		-16.66	-29.2	0.93 ± 0.55	1.53
Q353A		-24.13	-52.7	0.33 ± 0.01	1.10
T362A		-16.05	-28.8	2.10 ± 0.06	1.15
R365A		-18.97	-37.9	1.43 ± 0.11	1.13
R394A		-22.69	-47.2	0.26 ± 0.01	1.01
MEX-3C KH1	CAGAGU			ND	
	GUUUAG	-7.82	-6.15	32.57 ± 2.04	1.47
MEX-3C KH2	CAGAGU	-16.00	-29.4	3.15 ± 0.25	1.17
	GUUUAG	-16.36	-33.3	11.79 ± 0.45	1.35

Bold and italic fonts are used to highlight the mutant nucleotides of MRE10 RNA.