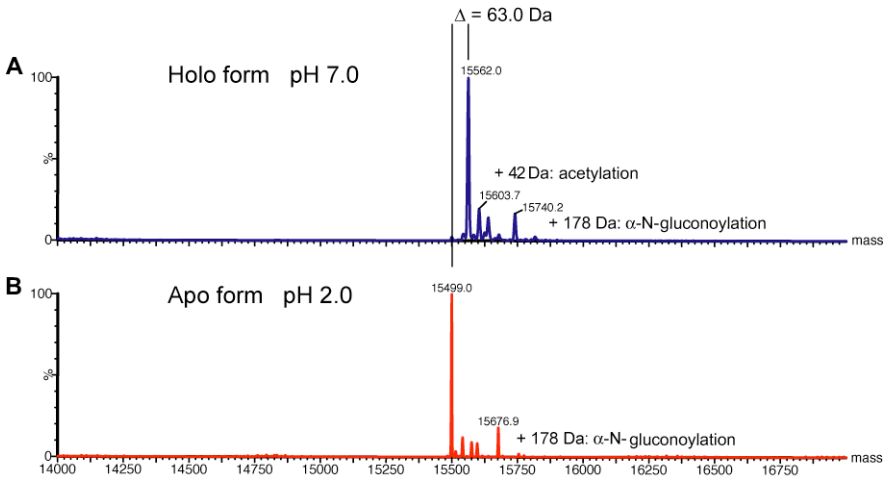
Supplementary Data

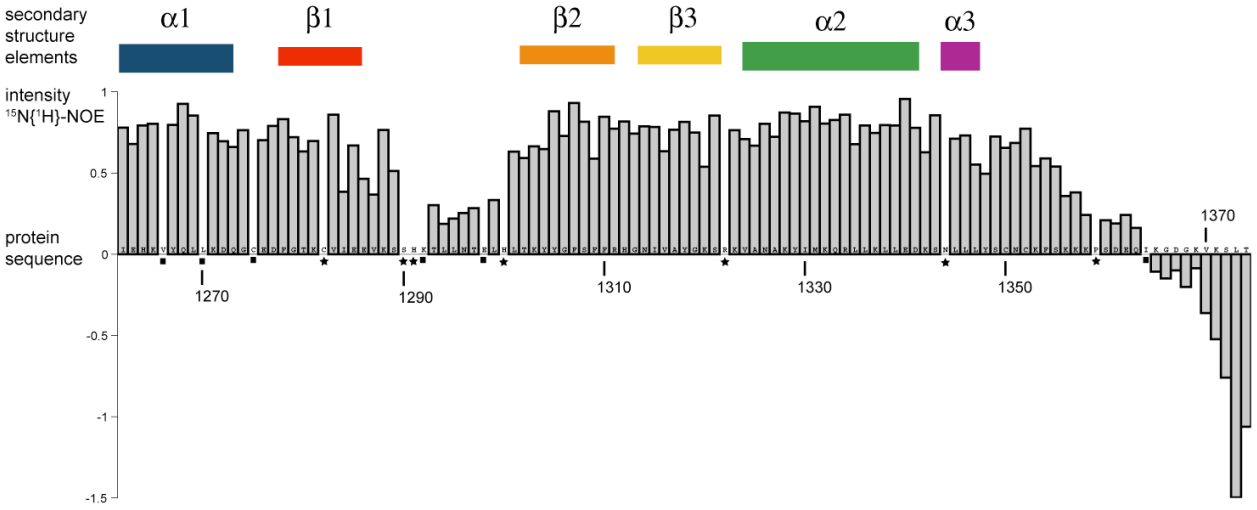
An Extended dsRBD with a Novel Zinc-Binding Motif Mediates Nuclear Retention of Fission Yeast Dicer

Pierre Barraud, Stephan Emmerth, Yukiko Shimada, Hans-Rudolf Hotz,
Frédéric H.-T. Allain, and Marc Bühler

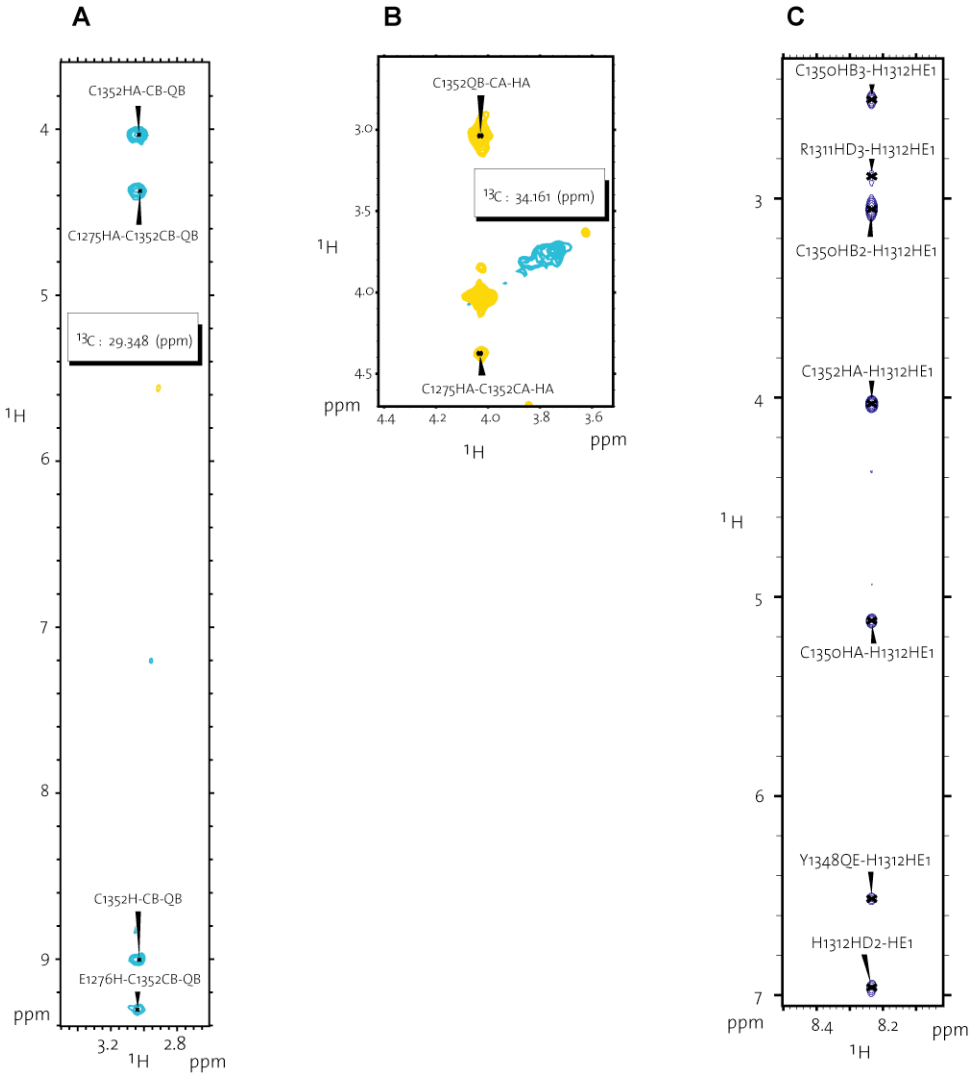
Supplementary Figures



Supplementary Figure 1.
The C-terminal domain of Dcr1 binds one zinc ion per protein domain.
(A) ESI-TOF deconvoluted spectrum obtained under non-denaturing conditions (10 mM ammonium acetate pH 7.0). (B) ESI-TOF deconvoluted spectrum obtained under denaturing conditions (50 % acetonitrile and 0.1 % formic acid pH 2.0). Minor peaks come from acetylation, α-N-gluconoylation of the histidine tag and phosphorylation.

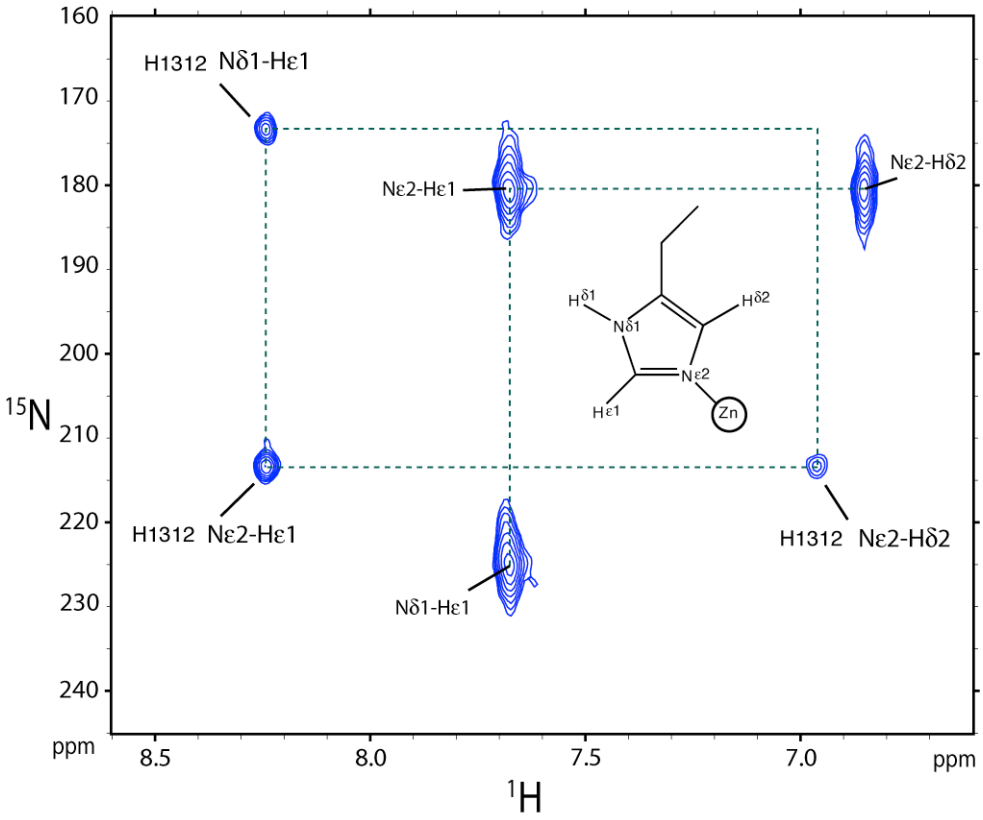


Supplementary Figure 2.
Residues of C33 following the canonical dsRBD fold are structured.
$^{15}N\{^{1}H\}$-NOE values plotted against the Dcr1 dsRBD sequence showing that the residues following the canonical dsRBD fold are structured. Values for residues marked with black squares could not be determined due to resonance overlap. Residues marked with an asterisk correspond to the one for which the amide assignment is missing. The secondary structure elements are represented as rectangles above the diagram. The color code refers to the one used in Figure 1.

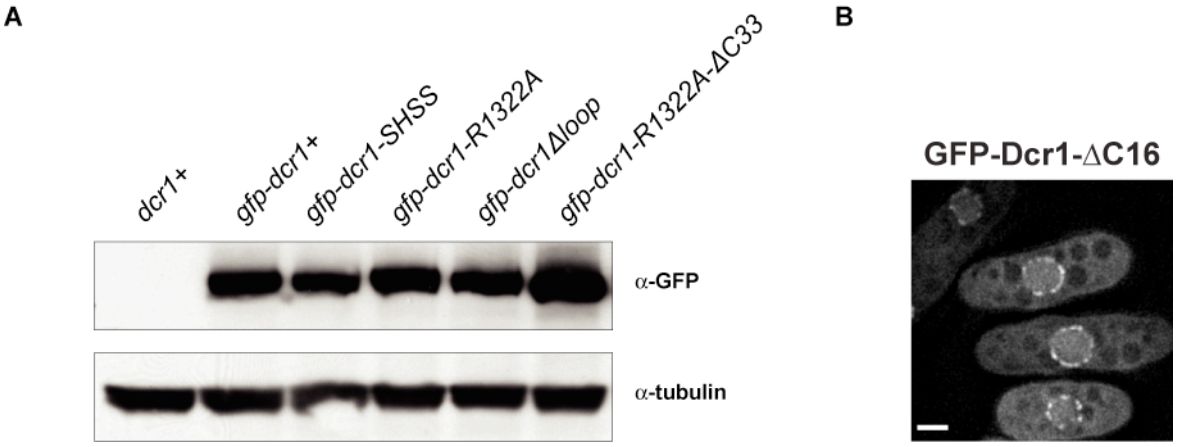


Supplementary Figure 3.

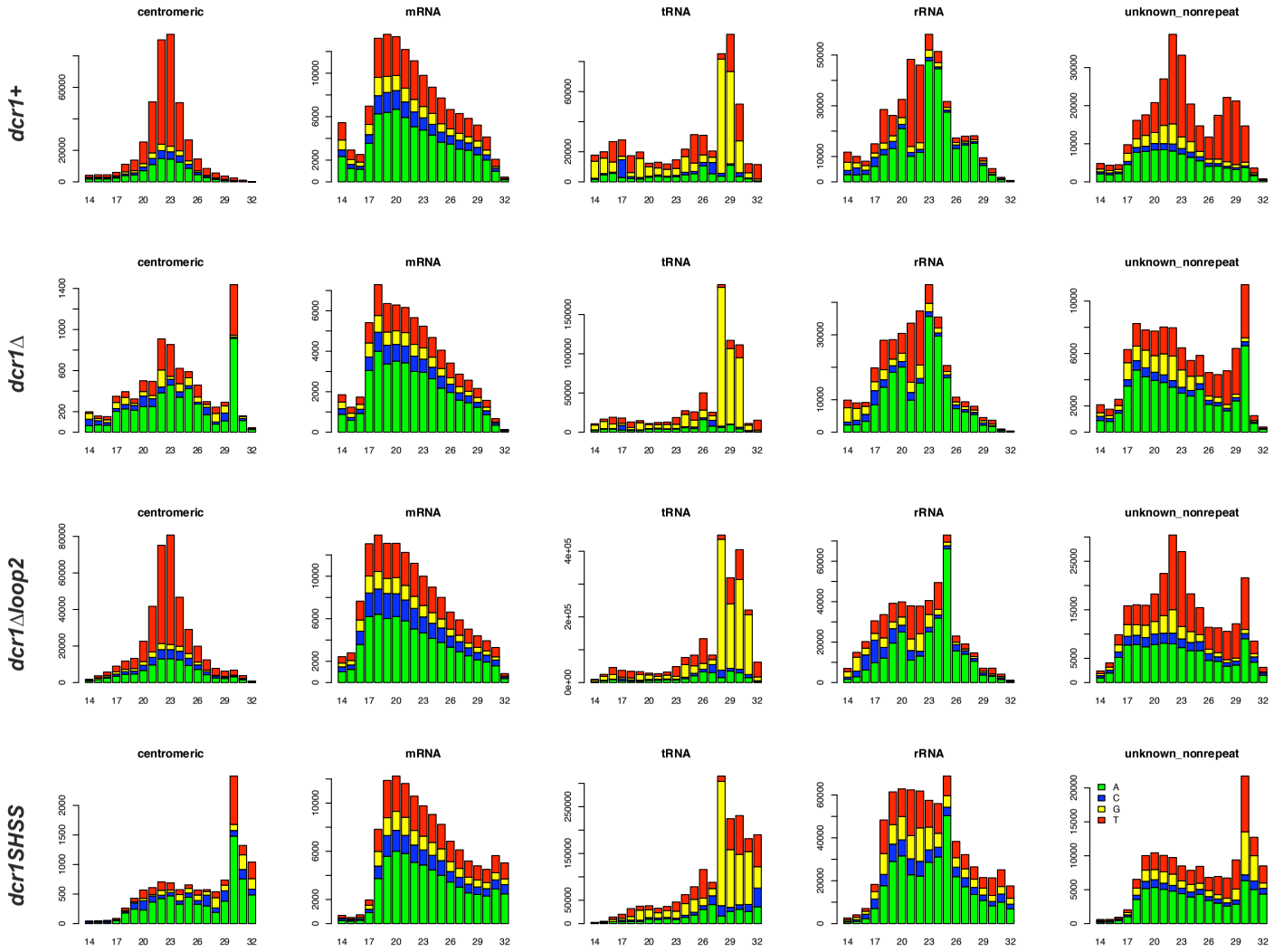
Coordination sphere around the zinc. (A-C) NOE contacts supporting the coordination sphere around the zinc ion. (A) Slice of the aliphatic 3D NOESY-(^{1}H-^{13}C)-HSQC showing contacts between C1352 and C1275. (B) Slice of the aliphatic 3D NOESY-(^{1}H-^{13}C)-HSQC showing contacts between C1352 and C1275. (C) Region of the 2D (^{1}H-^{1}H)-NOESY showing contacts between H1312 Hε1 and C1350 and C1352.



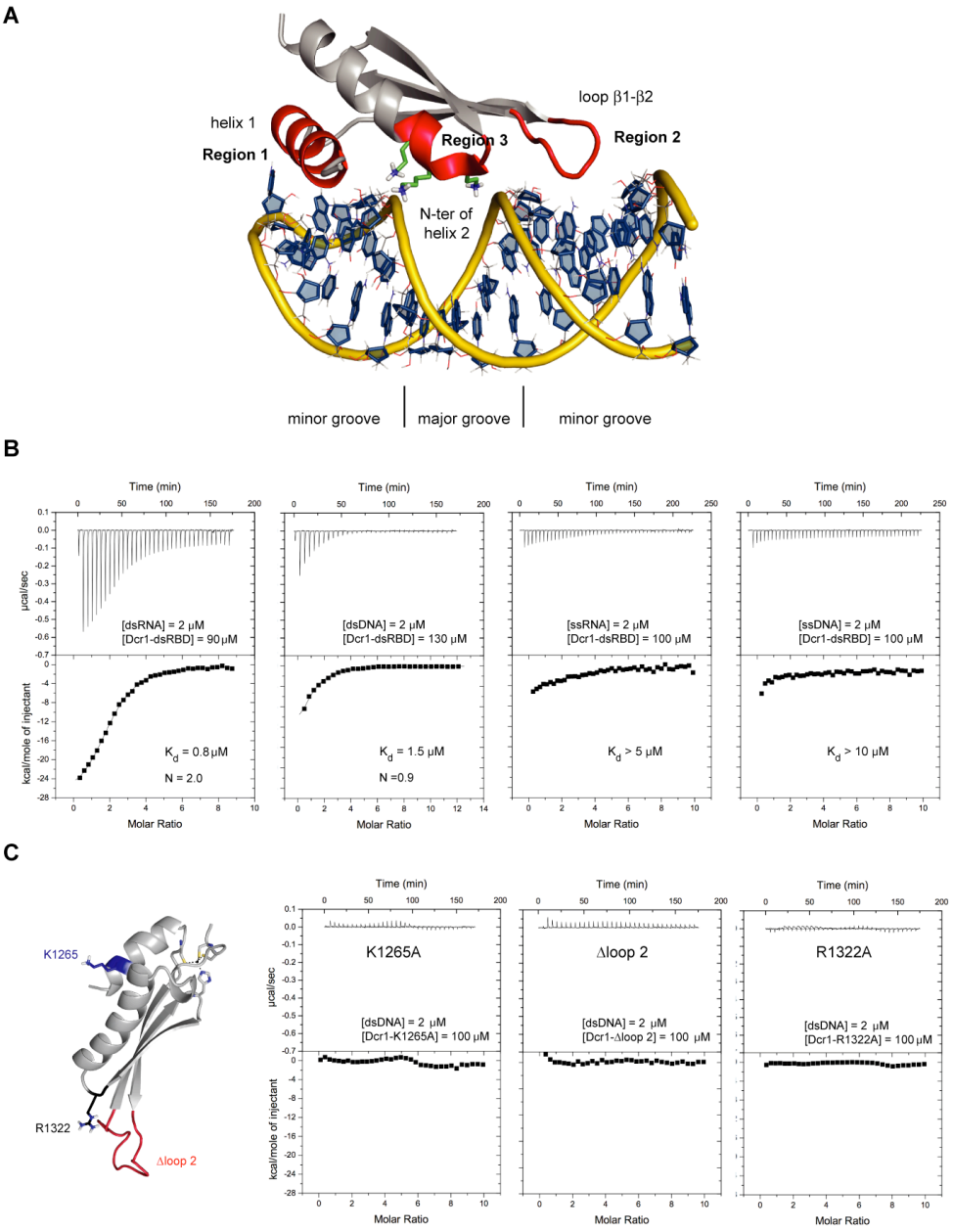
Supplementary Figure 4.
Long-range ^{1}H-^{15}N HSQC spectrum for the determination of the coordination mode of histidine 1312.
The pattern of cross peaks for H1312 is uniquely compatible with protonation on Nδ1 and therefore coordination through the Nε2 atom (c.f. Legge et al. (2004) J. Mol. Biol. 343, 1081-1093). The other peaks correspond to histidines from the N-terminal 6-histidine tag and to histidines present in the unstructured loop 2 (See Figure 1B). The pattern is characteristic of the histidine tautomer protonated on the Nε2.



Supplementary Figure 5.
Western blot analysis and imaging of GFP-Dcr1-ΔC16. (A) Western blot showing that stability is not affected for the zinc coordination motif mutant or the RNA-binding mutants. (B) Live-cell imaging of GFP-Dcr1-ΔC16. Scale bars = 2 μm.

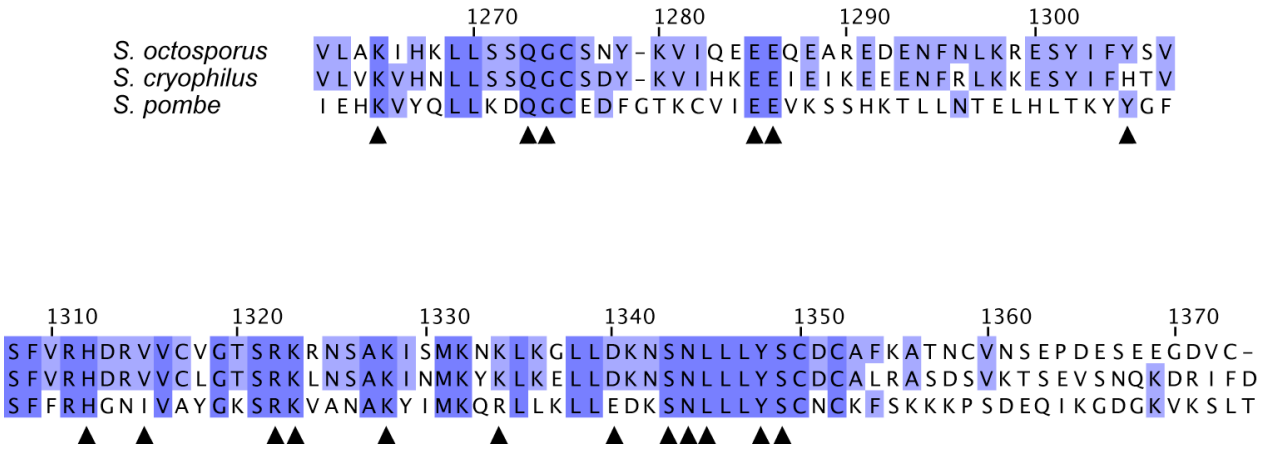


Supplementary Figure 6.
Size distribution and the 5'-most nucleotide of major classes of sRNAs sequenced from *dcr1+*, *dcr1Δ, dcr1Δloop2,* and *dcr1-SHSS* cells.

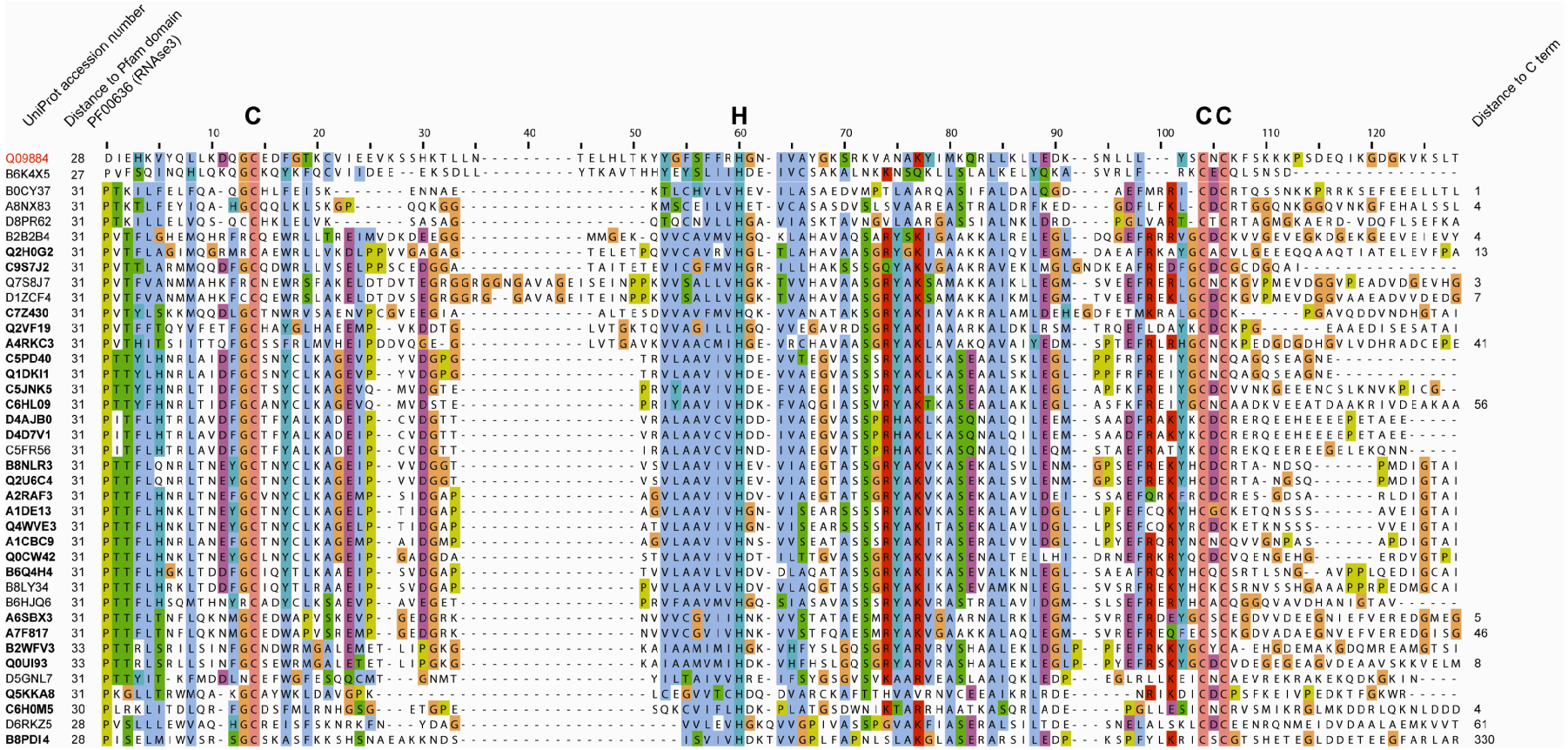


Supplementary Figure 7.

The Dcr1 dsRBD binds preferentially to dsRNA. (A) Regions of interaction between a classical dsRBD and a dsRNA target shown on the structure of ADAR2 dsRBD1 in complex with RNA (PDB code 2L3C). The three regions of interactions are represented in red on the protein structure (helix 1 and loop β1-β2 interact with the minor groove at one turn of interval and the N-terminal part of helix 2 interacts across the major groove). (B) ITC measurements of *S. pombe* Dcr1 dsRBD with different nucleic acid targets. From left to right: dsRNA (24 bp), dsDNA (24 bp), ssRNA (24 nt) and ssDNA (24 nt). (C) ITC measurements with dsDNA and Dcr1 dsRBD mutants (R1322A, Δloop 2 and K1265A).



Supplementary Figure 8.
Identification of the solvent exposed and conserved residues of Dcr1 dsRBD.
Black arrows show residues with a solvent exposed side chain and conserved biochemical properties among *S. pombe*, *S. octosporus* and *S. cryophilus*. Those residues are displayed in color on Figure 5A. Residue numbers refer to the *S. pombe* Dcr1 protein (UniprotKB Q09884).



Supplementary Figure 9.

Conservation of CHCC motif in C-termini of RNase III enzymes. Alignment of all annotated RNase III proteins with C termini that contain the $_{CHCC}$ motif. UniProt accession for the *S. pombe* Dcr1 protein is highlighted in red. Accessions in bold are proteins expressed in human or plant pathogenic yeast as listed in Supplementary Table II.

Supplementary Table I

Small RNAs sequenced from wild-type and Dcr1 mutant cells

	dcr1+		**dcr1Δ**		***dcr1-SHSS***		***dcr1Δloop2***	
	reads	%	reads	%	reads	%	reads	%
rRNA+	437259	23	345175	27	693100	25	499969	14
rRNA-	1122	0.1	158	0	739	0	451	0
tRNA+	544340	29	719518	56	1752112	63	2155557	61
tRNA-	460	0	244	0	365	0	226	0
snRNA+	5728	0.3	7382	0.6	9529	0.3	5786	0.2
snRNA-	26	0	4	0	12	0	22	0
snoRNA+	2074	0.1	1165	0.1	2938	0.1	2136	0.1
snoRNA-	201	0	112	0	338	0	326	0
3UTR+	14172	0.7	6866	0.5	10732	0.4	11557	0.3
3UTR-	1312	0.1	533	0	439	0	1042	0
5UTR+	2835	0.1	1301	0.1	2269	0.1	2627	0.1
5UTR-	163	0	56	0	138	0	194	0
mRNA+	126176	6.6	64120	5	117254	4.2	135565	3.8
mRNA-	12114	0.6	5243	0.4	6013	0.2	10885	0.3
intron+	10031	0.5	7536	0.6	10126	0.4	15736	0.4
intron-	465	0	324	0	1006	0	432	0
centromeric+	401625	21	7123	0.6	9712	0.3	382979	11
centromeric-	15473	0.8	1357	0.1	2286	0.1	15895	0.4
telomeric+	1	0	0	0	0	0	2	0
telomeric-	3	0	1	0	1	0	0	0
mating_type_region+	798	0	8	0	2	0	756	0
mating_type_region-	1315	0.1	17	0	7	0	1292	0
rep_origin+	56	0	21	0	20	0	59	0
rep_origin-	5	0	1	0	2	0	12	0
pseudogene+	316	0	35	0	89	0	315	0
pseudogene-	261	0	14	0	21	0	204	0
LTR+	138	0	75	0	81	0	104	0
LTR-	55	0	45	0	30	0	72	0
wtf+	450	0	298	0	435	0	539	0
wtf-	6	0	4	0	5	0	5	0
misc_RNA+	18258	1	11769	0.9	19463	0.7	23950	0.7
misc_RNA-	5558	0.3	861	0.1	1255	0	3151	0.1
unknown_non-repeat	303920	16	103173	8	148691	5.3	274583	7.7
sum	1906716	100	1284539	100	2789210	100	3546429	100

Weighted number of small RNA reads for different genomic features from deep sequencing of total sRNAs in wild-type, *dcr1Δ, dcr1-SHSS, and dcr1Δloop2* mutant cells.

Supplementary Table II

Pathogenic yeasts expressing at least one dicer protein containing the CHCC motif

Accession	Organism	Pathology
Q5KKA8	**Cryptococcus neoformans**	Cryptococcosis (opportunitstic infection in HIV+ patients, fungal meningitis)
Q2H0G2	Chaetomium globosum	Peritonitis, Onychomycosis
C5PD40	Coccidioides posadasii	Coccidioidomycosis (California Disease, Valley Fever)
Q1DKI1	Coccidioides immitis	Coccidioidomycosis (California Disease, Valley Fever)
C5JNK5	**Ajellomyces dermatitidis**	Blastomycosis (Chicago disease, Gilchrist's disease)
C6HL09	Ajellomyces capsulata	Darling's Disease
D4AJB0	Arthroderma benhamiae	Dermatophytosis (Ringworm)
B8NLR3	Aspergillus flavus	Aspergillosis (lung, occasionally identified as the cause of corneal, otomycotic and nasoorbital infections)
Q2U6C4	Aspergillus oryzae	Aspergillosis (only occasionlly pathogenic though)
A2RAF3	Aspergillus niger	Aspergillosis
Q4WVE3	**Aspergillus fumigatus**	Major cause of Aspergillosis (both invasive and allergic Aspergilloses), Mycotoxicosis, Farmer's Lung
A1DE13	Neosartorya fischeri	Aspergillosis
A1CBC9	Aspergillus clavatus	Aspergillosis, Mycotoxicosis
Q0CW42	**Aspergillus terreus**	Aspergillosis of the lungs and or disseminated aspergillosis, Onychomycosis, Mycotoxicosis (thermotolerant yeast)
B6Q4H4	Penicillium marneffei	Penicilliosis (third most common opportunistic infection in HIV+ patients)
D4D7V1	Trichophyton verrucosum	Dermatophytosis (Ringworm)
C7Z430	Nectria haematococca	Plant diseases, opportunistic infections in animals
C9S7J2	Verticillium albo-atrum	Verticillium Wilt Disease
Q2VF19	Cryphonectria parasitica	Chestnut Blight
A4RKC3	**Magnaporthe grisea**	Blast Disease (rice)
A6SBX3	Botryotinia fuckeliana	Gray mold disease
A7F817	Sclerotinia sclerotiorum	White Mold
B2WFV3	Pyrenophora tritici-repentis	Tan spot
Q0UI93	Phaeosphaeria nodorum	Glume blotch (major pathogen of wheat)
C6H0M5	Mucor c. f. lusitanicus	plant diseases
B8PDI4	Postia placenta	destructive decay of wood in buildings and other structures

UniProt release 2010_10 - Oct 5, 2010. Gold, human or animal pathogens. Yellow, plant pathogens.

Strains used in this study

Strain	Genotype	Source
SPB80	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+*	*2*
SPB81	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::TAP-Kan*	*2*
SPB287	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-ΔC33::Nat*	*1*
SPB278	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1+*	*1*
SBP777	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-SHSS::Nat*	*1*
SPB776	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-SHSS::Nat*	*1*
SPB773	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-ΔC16::Nat*	*1*
SBP771	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-Δloop2::Nat*	*1*
SPB770	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-Δloop2::Nat*	*1*
SBP773	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-ΔC16::Nat*	*1*
SPB772	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-ΔC16::Nat*	*1*
SPB919	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-R1322A::Hph*	*1*
SPB918	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr-R1322A::Hph*	*1*
SPB919	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-R1322A-ΔC33::Hph*	*1*
SPB949	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr-K1265A::Hph*	*1*
SPB775	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-Y1348A::Nat*	*1*
SPB1079	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-R1334A::Nat*	*1*
SPB1080	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-S1349A::Nat*	*1*
SPB1082	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-N1344A::Nat*	*1*
SPB1116	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-Y1348A-N1344A-S1349A::Nat*	*1*

1 = Bühler lab strain collection, 2 = obtained from Danesh Moazed