

Spliceosome:

The crucial role of Prp8

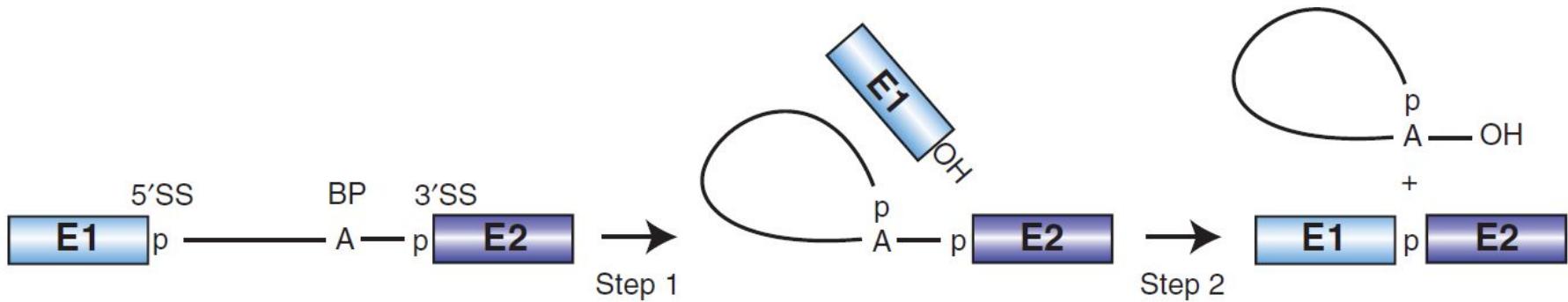
Structural Biology

Ariadna Martínez, Laura Sans, Marta Vilademunt and Mar Villa

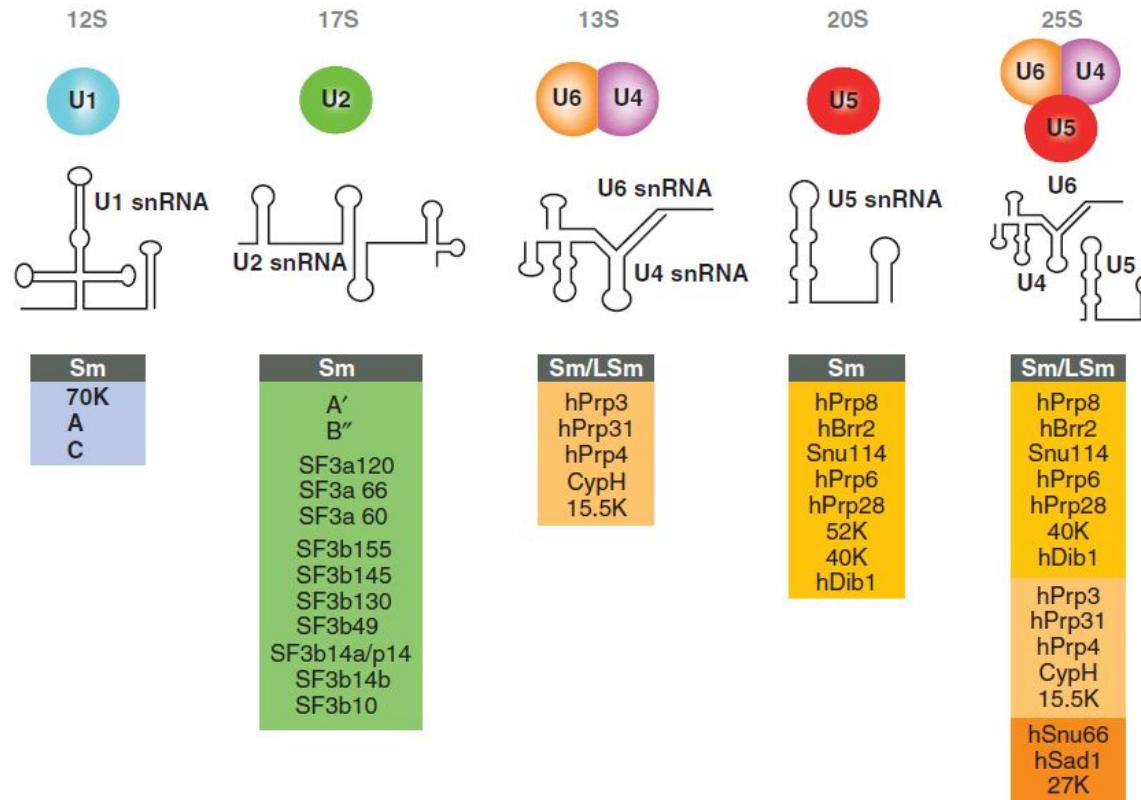
Index

- 1. Spliceosome basics**
 - a. Introduction - Branching and exon ligation
 - b. Assembly and activation of spliceosome -
Steps
 - c. Splicing active site
- 2. Prp8 analysis**
 - a. N-terminal (N-domain)
 - b. Reverse transcriptase (RT) fingers/palm
 - c. Thumb/X
 - d. Linker
 - e. RNase H-like
 - f. Jab1/MPN
- 3. Aar2**
- 4. Transition Aar2-Brr2**
- 5. Brr2**

Spliceosome basics - Branching and exon ligation

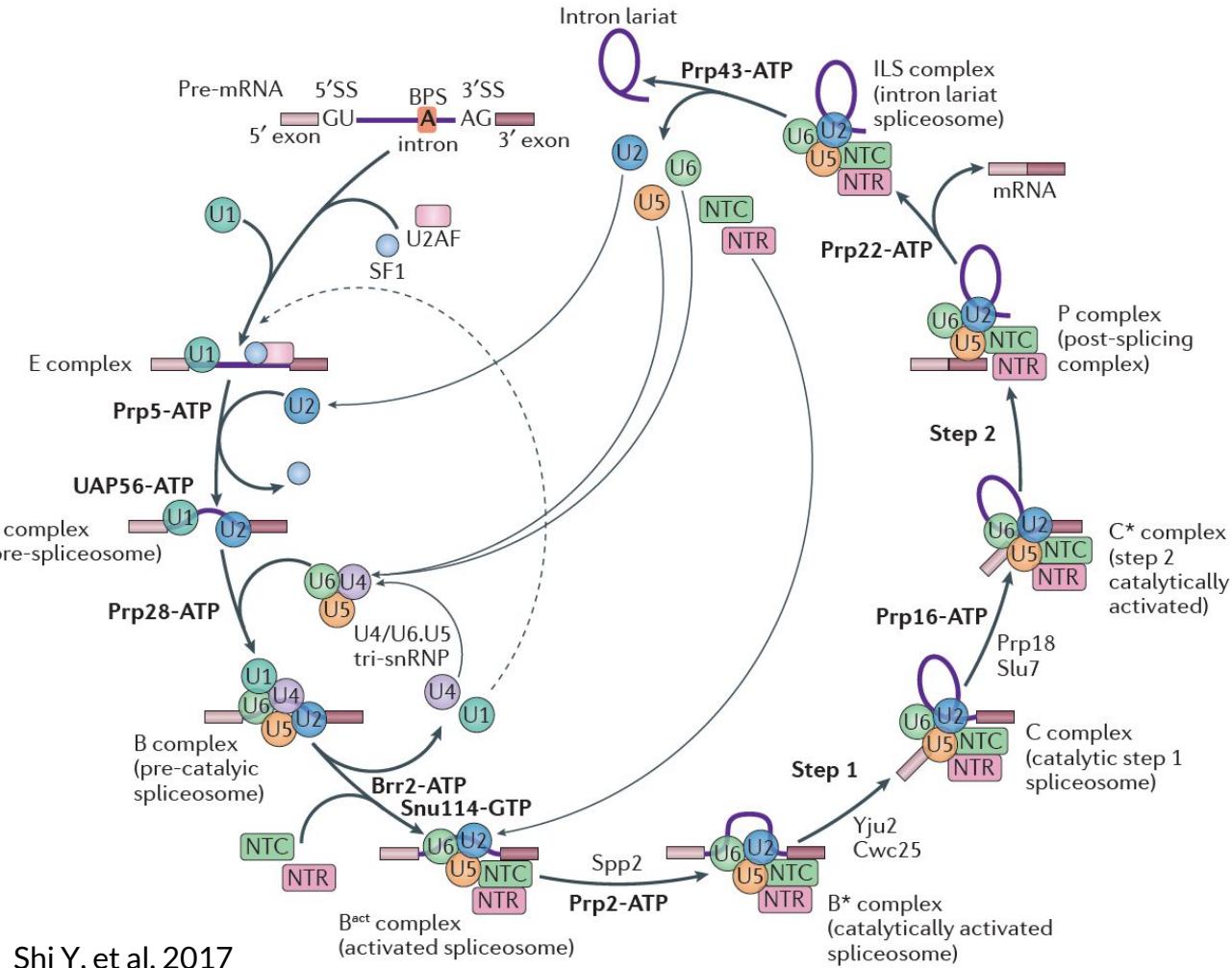


Spliceosome basics - Branching and exon ligation

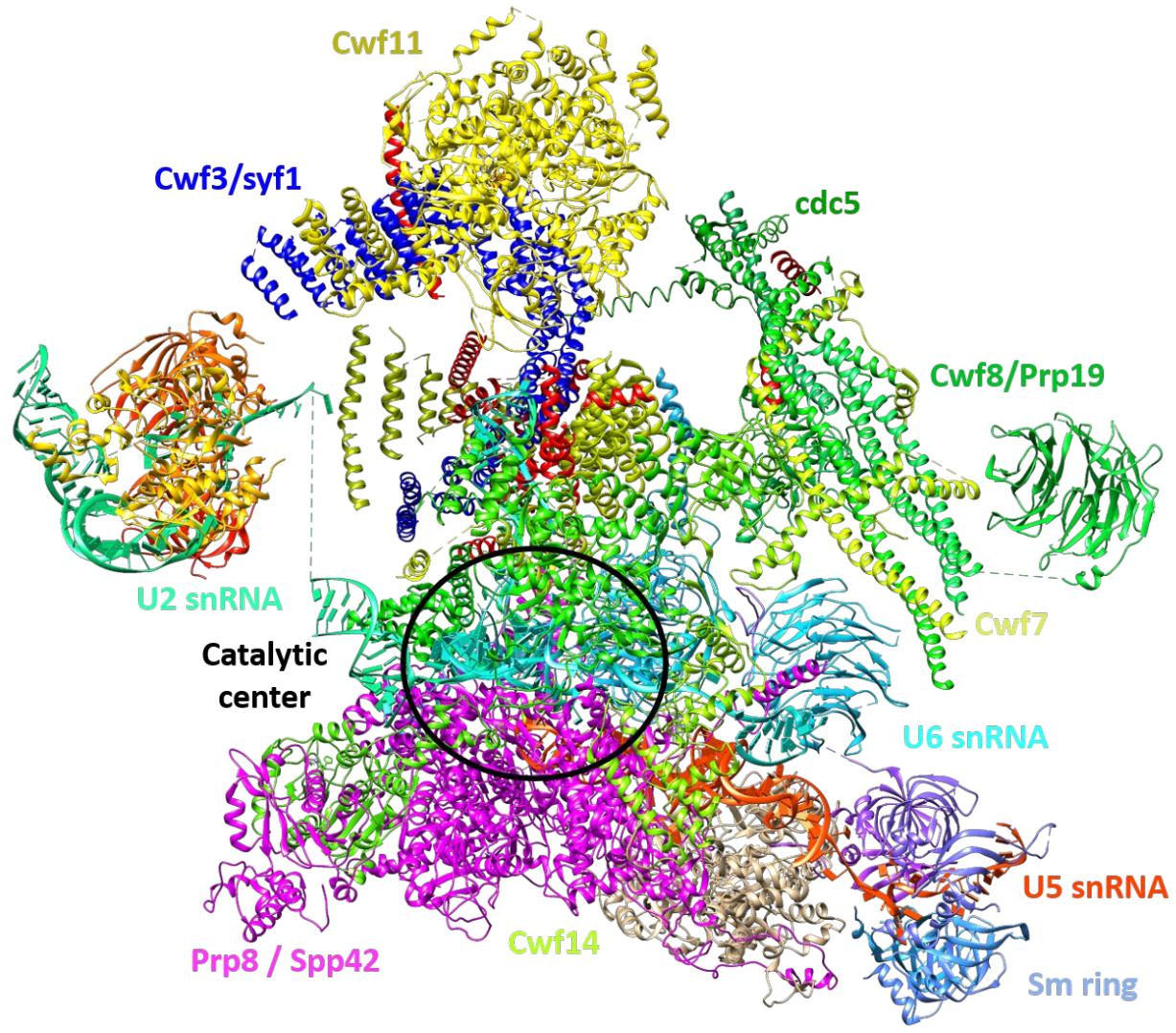


Will CL, et al. 2011

Spliceosome basics - STEPS

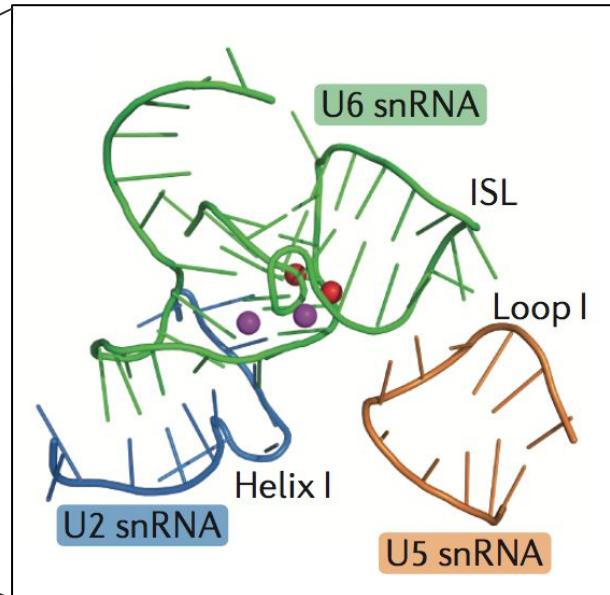
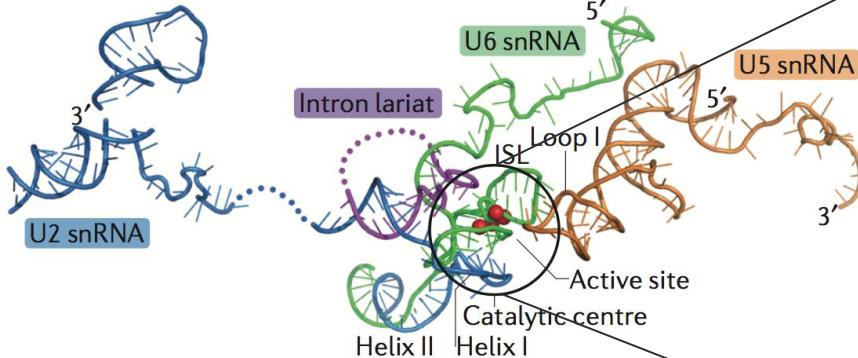


Structure of yeast spliceosome



PDB code: 3JB9
Resolution: 3,6 Å

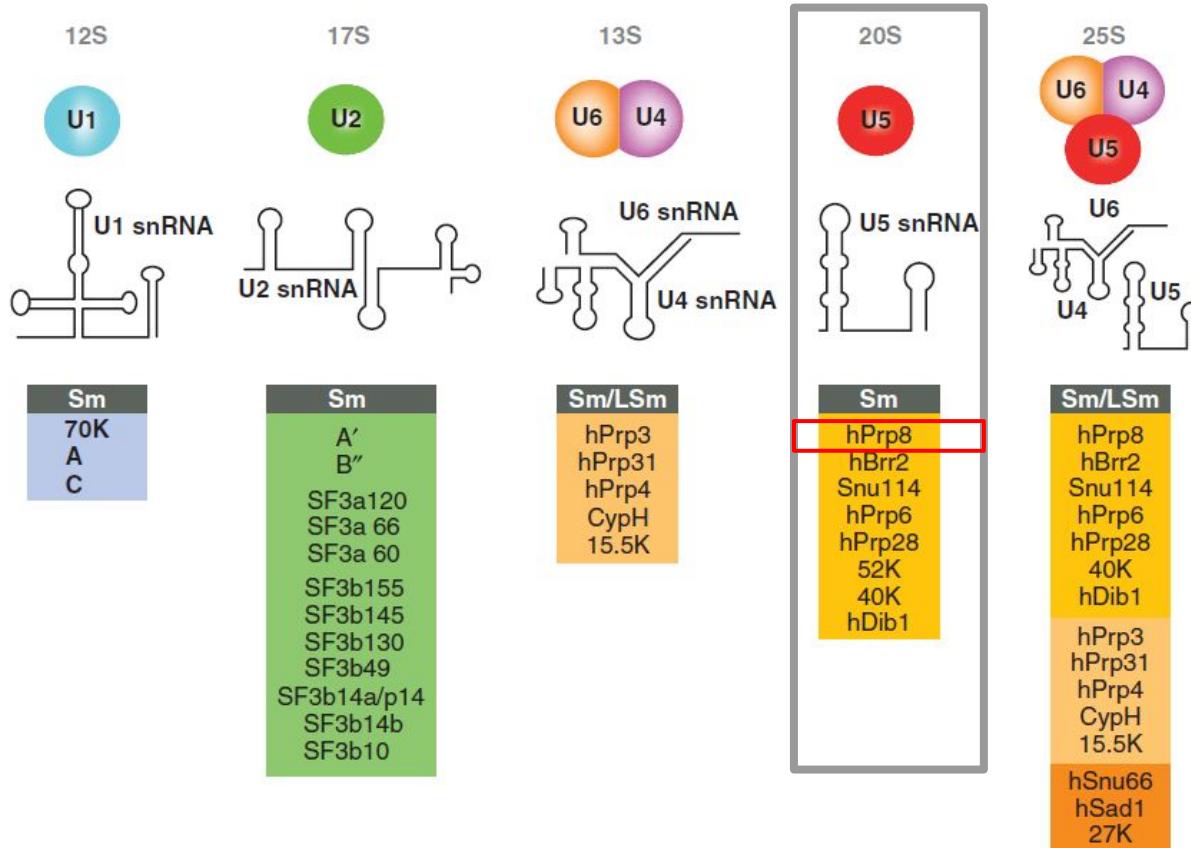
Catalytic core



- Intramolecular stem-loop (ISL) of U6 snRNA
- helix I of the U2-U6 duplex
- Mg^{2+} ions
- loop I of U5 snRNA

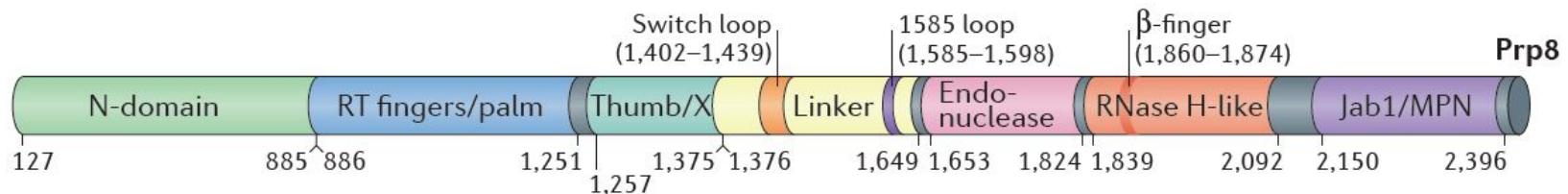
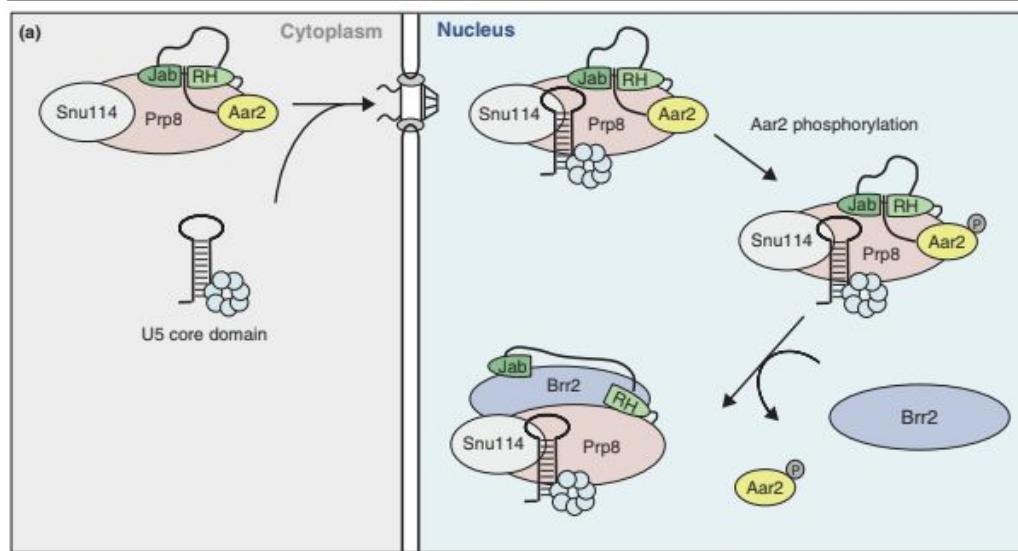
Shi Y, et al. 2017

U5 snRNP



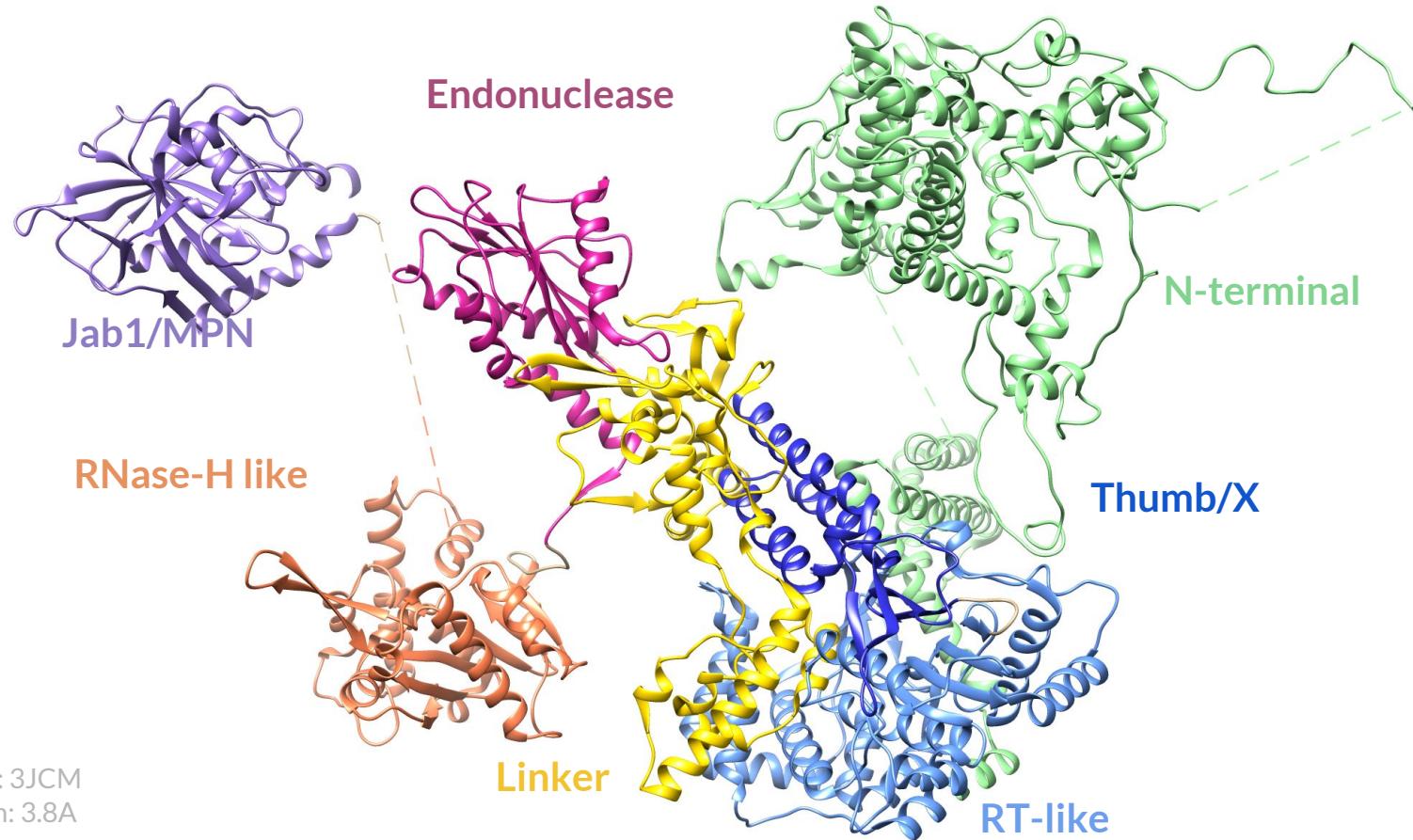
Will CL, et al. 2011

Prp8



Shi Y, et al. 2017

Prp8



PDB code: 3JCM
Resolution: 3.8A

Prp8

Class	Alpha and beta proteins
Fold	Ribonuclease H-like motif (3 layers: <i>a/b/a</i> ; mixed beta-sheet of 5 strands, order 32145; strand 2 is antiparallel to the rest)
Superfamily	Ribonuclease H-like
Family	Prp8 beta-finger domain-like
Protein	Pre-mRNA-splicing factor 8
Species	<i>Saccharomyces cerevisiae</i>



Prp8

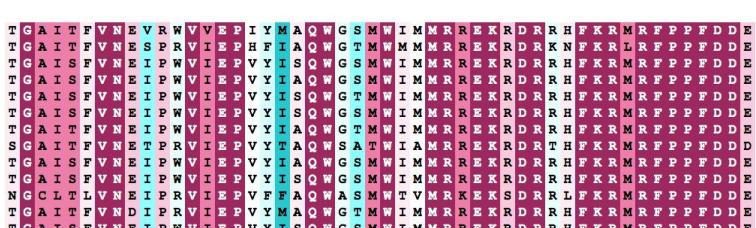
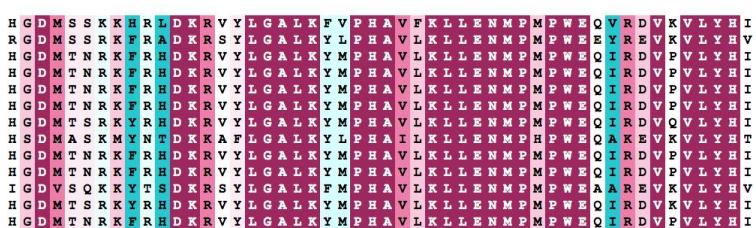
HMMalign from PFAM (of each domain)

	1261	1320
H.sapiens	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	001 A.thaliana
M.mulatta	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	002 S.pombe
C.lupus	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	003 C.lupus
M.musculus	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	004 X.tropicalis
G.gallus	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	005 D.rerio
D.rerio	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	006 M.musculus
D.melanogaster	msgfecrilm kcrtnqneefht HRDGWNLQN EITKERTAQ C FLRVVDESMQ RFHNVRQIL	007 D.melanogaster
C.elegans	msgfecrilm kcrtnaneefv HRDGWNLQN EITKERTAQ C FLKVDEESLS FKHNRRQIL	008 S.cerevisiae
S.cerevisiae	smcgfevrl prgrmeevs NDEGVWDLVD ERTKERTAQ C L YKVSEEIKL KFDSRIRGIL	009 G.gallus
S.pombe	mtgefvrilp kirqneefsl . KDGWNLTD NRTKERTAQ C FIRVTEGDIN QFGNRRQIL	010 H.sapiens
N.crassa	cgfevrlpk irqnqdefpv . KDSWVSLVD NTTKERTAHA FLOVTEEDIQ KFNNRRQIL	011 N.crassa
A.thaliana	cgfevrlpk irqnqgeafss T RDGVWNLQN EQTKERTAVA FLRADDEHMK VFENRVRQIL	012 C.elegans
X.tropicalis	mcgefcrilp kcrtsyeefht HKDGWNLQN EVTERTAQ C FLRVVDESMQ RFHNVRQIL	013 M.mulatta

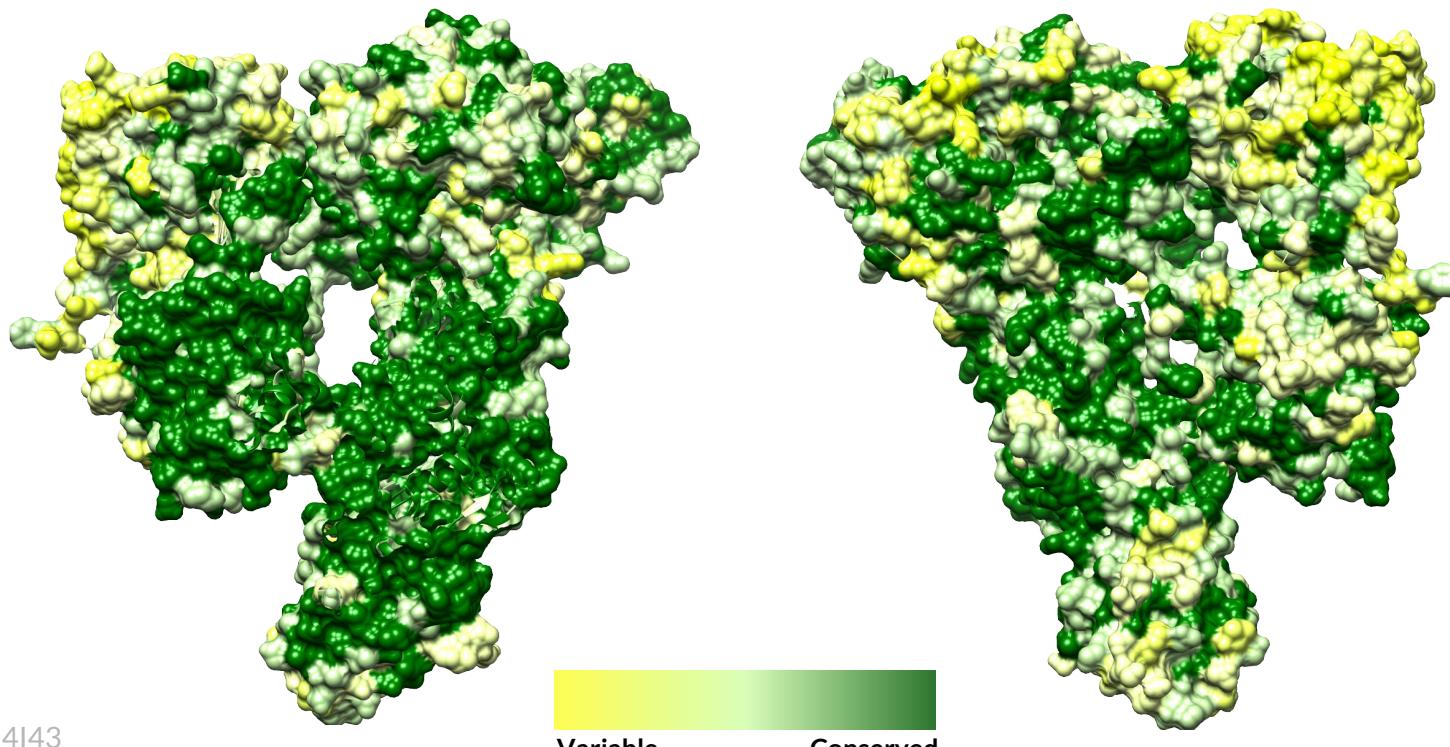
	1321	1380
H.sapiens	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	001 A.thaliana
M.mulatta	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	002 S.pombe
C.lupus	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	003 C.lupus
M.musculus	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	004 X.tropicalis
G.gallus	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	005 D.rerio
D.rerio	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	006 M.musculus
D.melanogaster	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	007 D.melanogaster
C.elegans	MSSGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	008 S.cerevisiae
S.cerevisiae	MASGSTTFK VAAKWNTSLI SLFTYFREAI VATEPDLIL VKGETRIONR VKGLGLNSKMP	009 G.gallus
S.pombe	MSSGSTTFK IANKWNTALI ALMTYREAA ISTPELL DLL VKCESKIQT R VKISLNSKMP	010 H.sapiens
N.crassa	MSSGSTTFK IANKWNTALI ALFTYREAA VSTVNLNTDI VKCETKIQT R VKIGLNSKMP	011 N.crassa
A.thaliana	MSSGSTTFK IVNKWNTALI GLMTYFREAT VHTQELLDLL VKCENKIQT R VKIGLNSKMP	012 C.elegans
X.tropicalis	MASGSTTFK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQT R IKIGLNSKMP	013 M.mulatta

	1381	1440
H.sapiens	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	001 A.thaliana
M.mulatta	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	002 S.pombe
C.lupus	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	003 C.lupus
M.musculus	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	004 X.tropicalis
G.gallus	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	005 D.rerio
D.rerio	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	006 M.musculus
D.melanogaster	SRFPPVVFY PKEGLGGM GLMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl	007 D.melanogaster
C.elegans	SRFPPVVFY PKEIGGLGGM GLMGHVLIPQS DLRW.mqte aggvthrgm shdedqlip	008 S.cerevisiae
S.cerevisiae	TRFPPAVFYT PKEGLGGM GLMGHVLIPQS DLSWSkqtdt githfragmt hedeklipti	009 G.gallus
S.pombe	SRFPPAVFYT PKEGLGGM GLMGHVLIPQS DLSWSkqtdt githfragmt hedeklipti	010 H.sapiens
N.crassa	SRFPPAVFYT PKEGLGGM GLMGHVLIPQS DLSWSkqtdt githfragmt hedeklipti	011 N.crassa
A.thaliana	SRFPPVIFT PKEIGGLGGM GLMGHVLIPQS DLRYSngtdv gvhfrsgms heedqlipnl	012 C.elegans
X.tropicalis	SRFPPVIFT PKEIGGLGGM GLMGHVLIPQS DLRYSngtdv gvhfrsgms heedqlipnl	013 M.mulatta

ConSurf Color Coded MSA

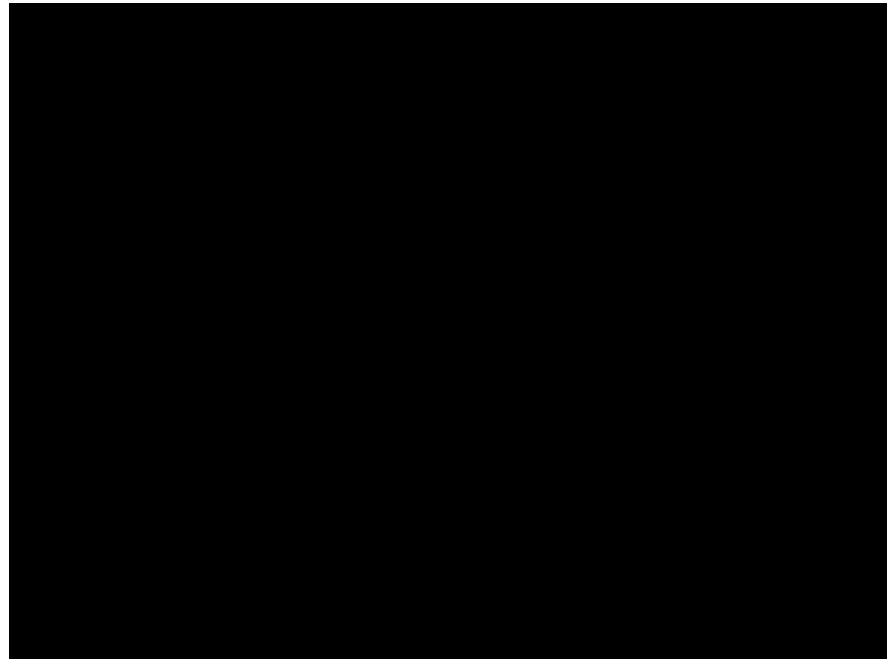
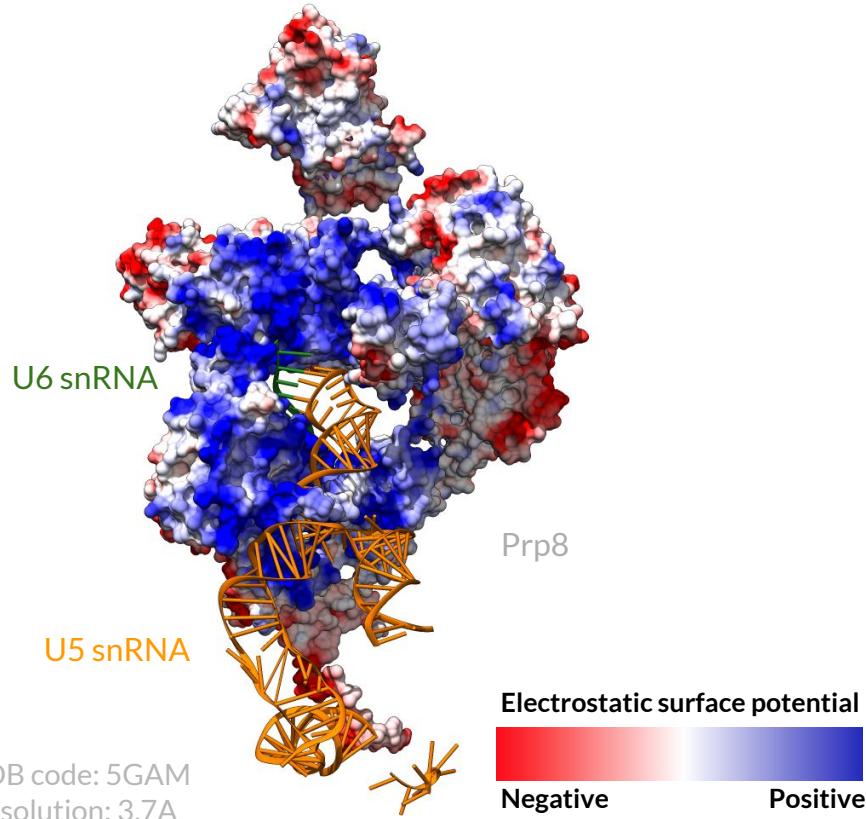


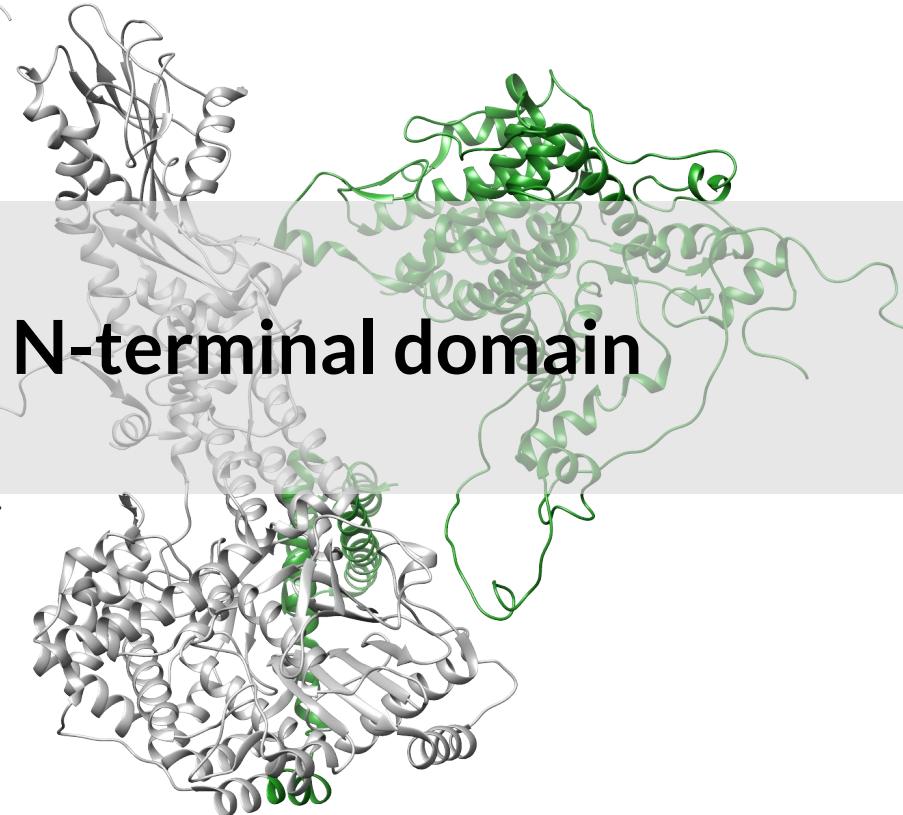
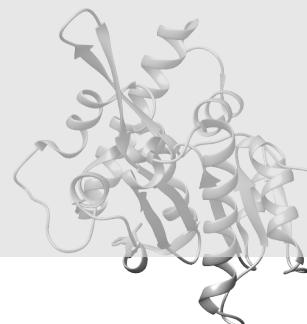
Prp8



PDB code: 4I43
Resolution: 2A

Electrostatic surface potential in the cavity





N-domain

RT fingers/palm

Thumb/X

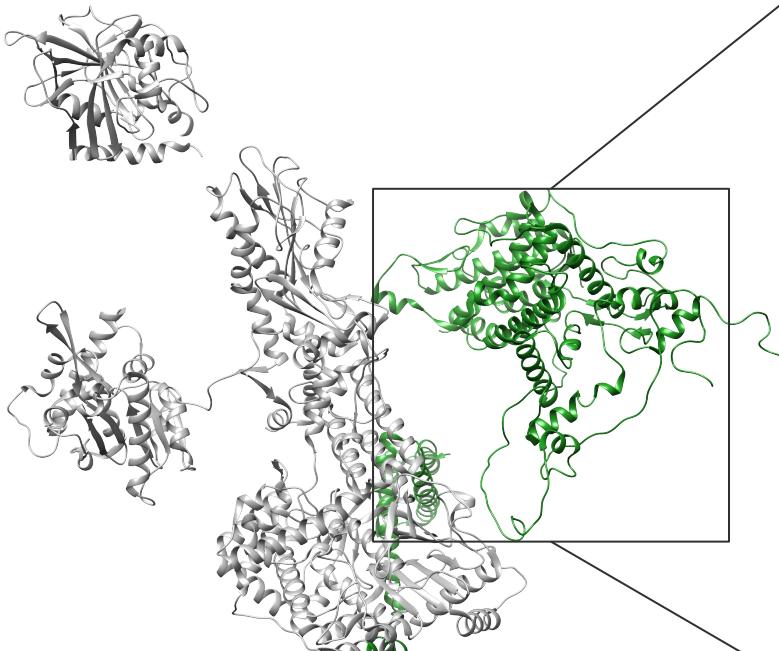
Linker

Endonuclease

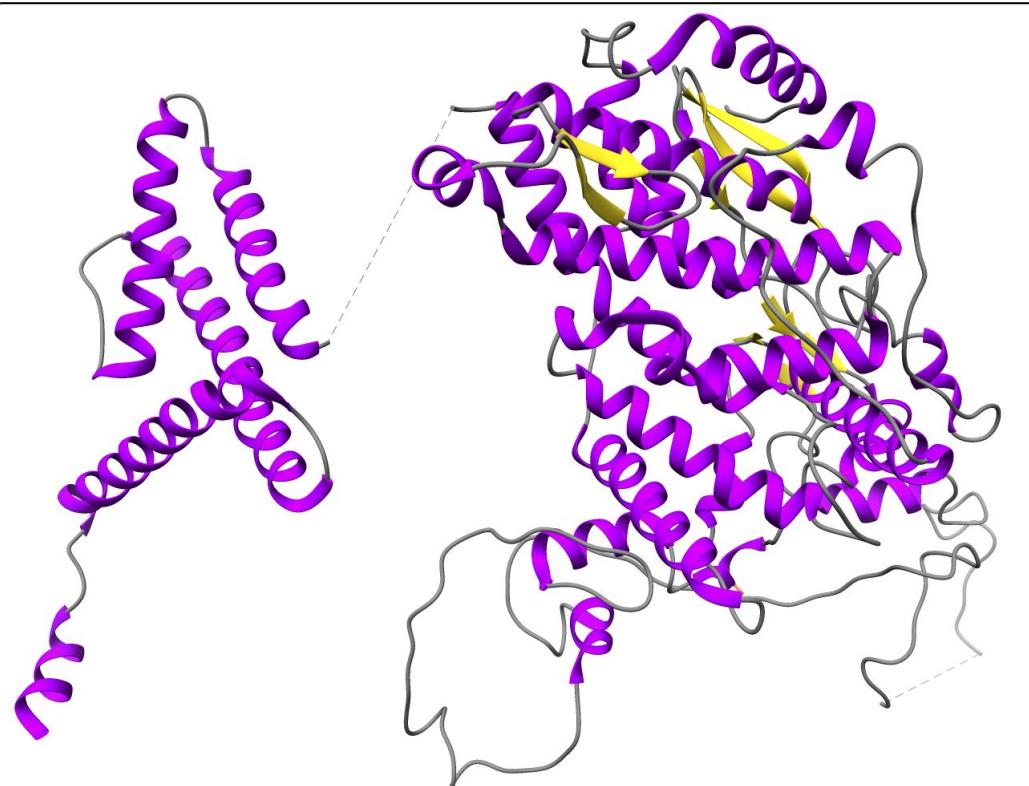
RNase H-like

Jab1/MPN

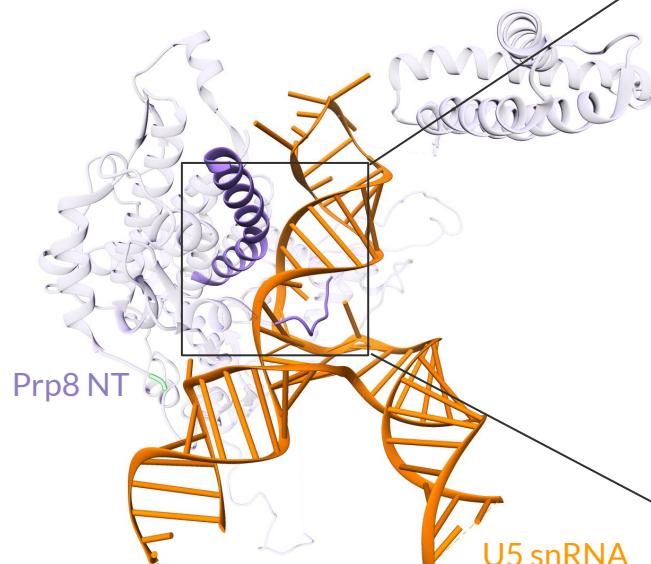
N-terminal



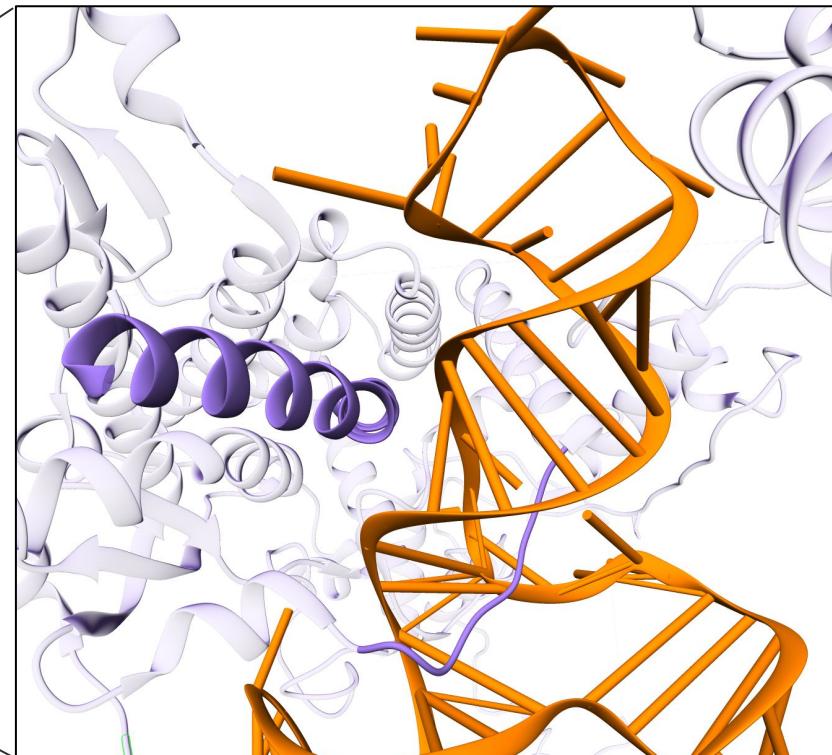
PDB code: 3JCM
Resolution: 3.8A



N-terminal



The N-domain recognises the U5 snRNA...

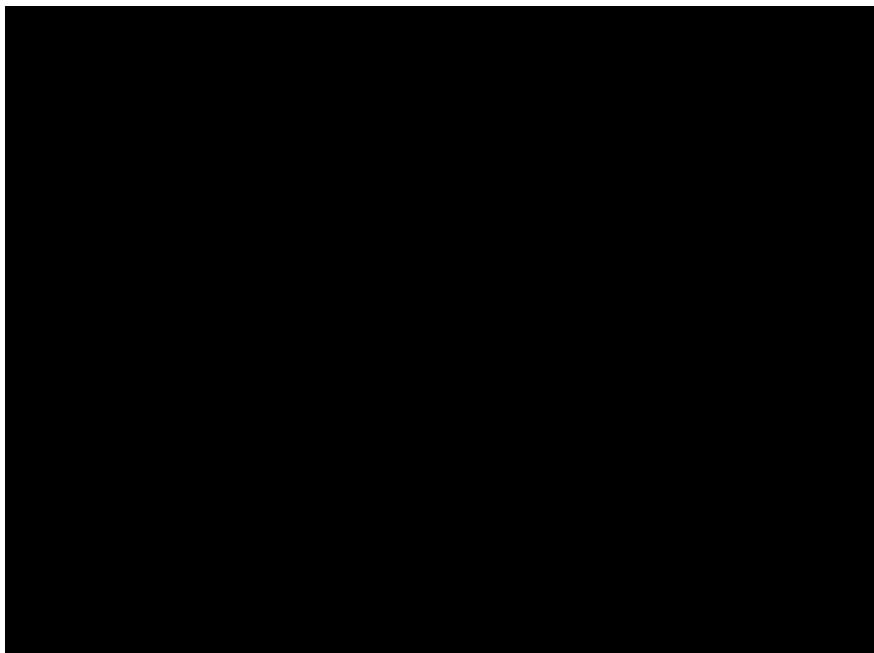
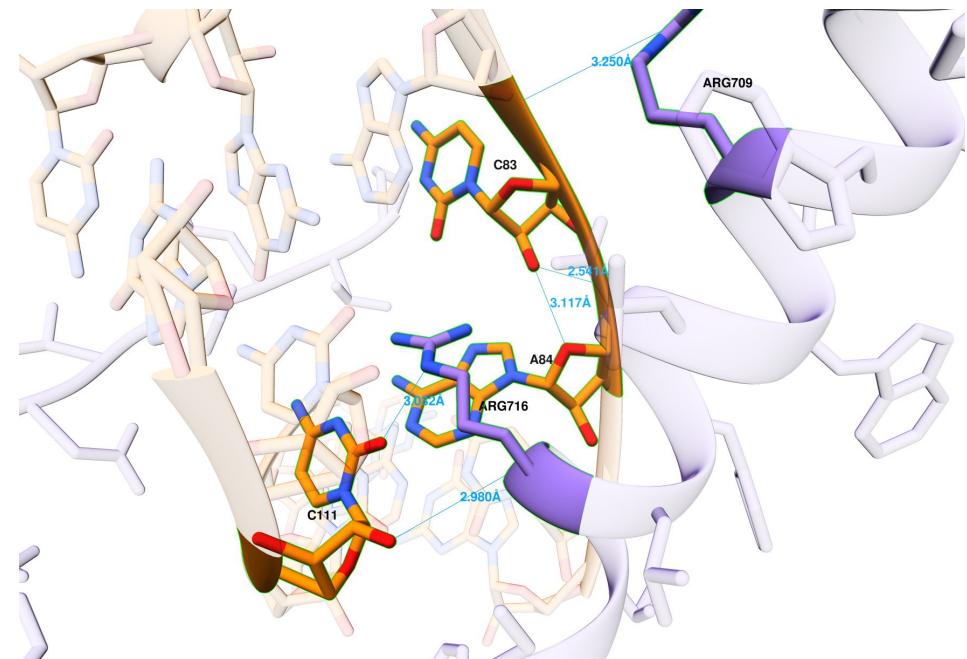


PDB code: 3JCM
Resolution: 3.8A

...and harnesses it with a polypeptide loop!

N-terminal

Closed-up view of the hidrogen bonds between the α -helix and the minor groove



N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

Jab1/MPN

N-terminal

 Negatively charged

 Positively charged

 Non-polar

 Polar

		661				720
H.sapiens	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
M.mulatta	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
C.lupus	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
M.musculus	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
G.gallus	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
D.rerio	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAPG	WRVWLFFMRG
D.melanogaster	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAPG	WRVWLFFMRG
C.elegans	HVGQLTGMYR	YKYKLMRQVR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAPG	WRVWLFFLRG
S.cerevisiae	HIGQLTGIYR	YKYKVMHQIY	ACKDLKHIIY	YFKFNK.NLJGK	GPGCGFWQPA	WRVWLNFRLG
S.pombe	HVGQLTGMYR	YKYRMLRQIR	ACKDFKHIIY	YRFNTgPVGK	GPGCGFWAPS	WRVWLFFLRG
N.crassa	HVGQLTGMYR	YKYKLMHQIR	SCKDLKHLIY	YRFNAgPVGK	GPGCGFWAPA	WRVWLFFMRG
A.thaliana	HVGQLTGMYR	YKYRMLRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAPM	WRVWLFFLRG
X.tropicalis	HVGQLTGMYR	YKYKLMRQIR	MCKDLKHLIY	YRFNTgPVGK	GPGCGFWAAG	WRVWLFFMRG
		721				780
H.sapiens	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
M.mulatta	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
C.lupus	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
M.musculus	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
G.gallus	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
D.rerio	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ
D.melanogaster	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRASVMHD	VDMMPEGIKQ
C.elegans	ITPLLERWLG	NLLSRQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPDGIKQ
S.cerevisiae	TIPLLERWYIG	NLITRQFEGR	.SNEIVKTTT	KQRLEDAYYDL	ELRNSVMDDI	LEMMPESIRQ
S.pombe	IVPLLERWLG	NLLARQFEGR	HSTGVAKQIT	KQRVDSHQDL	ELRAAVMNDI	LDIMIPEGIRQ
N.crassa	IIPLLERWLG	NLLSRQFEGR	HSKGVAKTVT	KQRVESHD	ELRASVMADL	LDMMPEGIKQ
A.thaliana	IVPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDV	VDAMPEGIKQ
X.tropicalis	ITPLLERWLG	NLLARQFEGR	HSKGVAKTVT	KQRVESHD	ELRAAVMHDI	LDMMPEGIKQ

Sequence alignment based on PFAM domains

N-domain

RT fingers/palm

Thumb/X

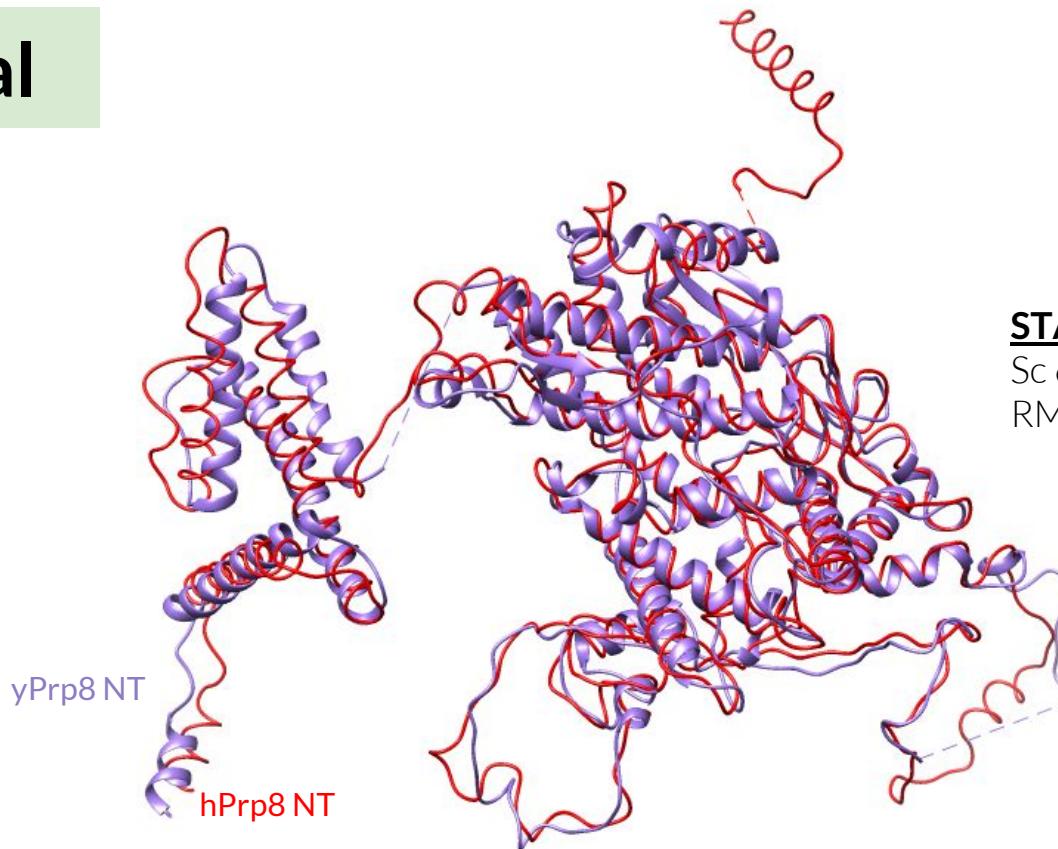
Linker

Endonuclease

RNase H-like

Jab1/MPN

N-terminal



STAMP VALUES

Sc 6.53
RMS 1.58

PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A

N-domain

RT fingers/palm

Thumb/X

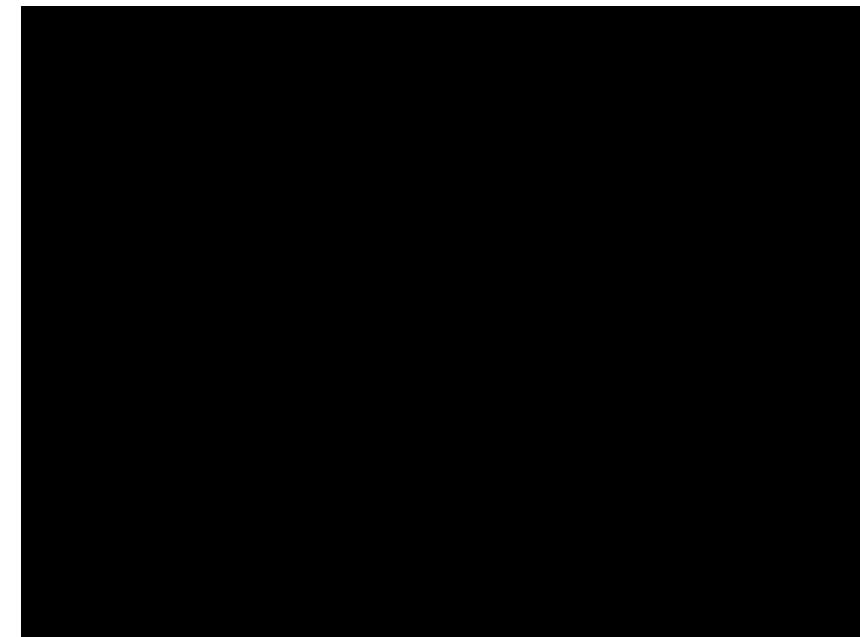
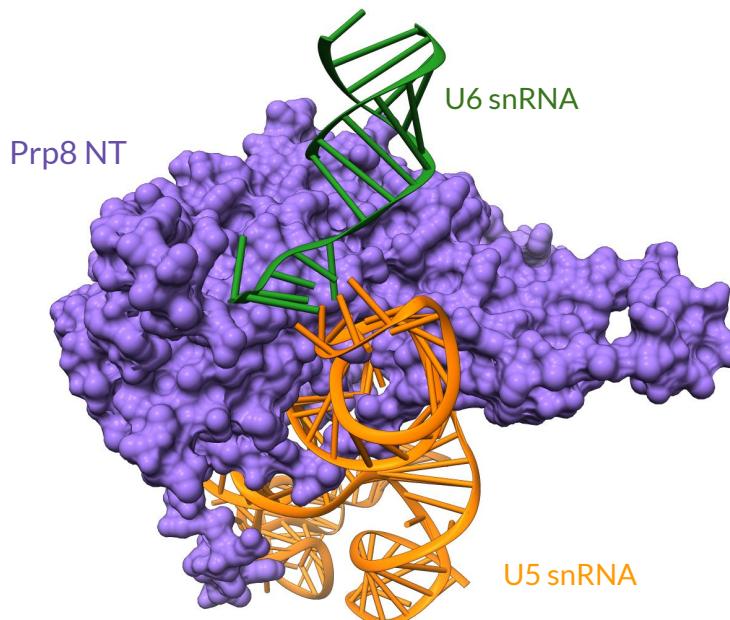
Linker

Endonuclease

RNase H-like

Jab1/MPN

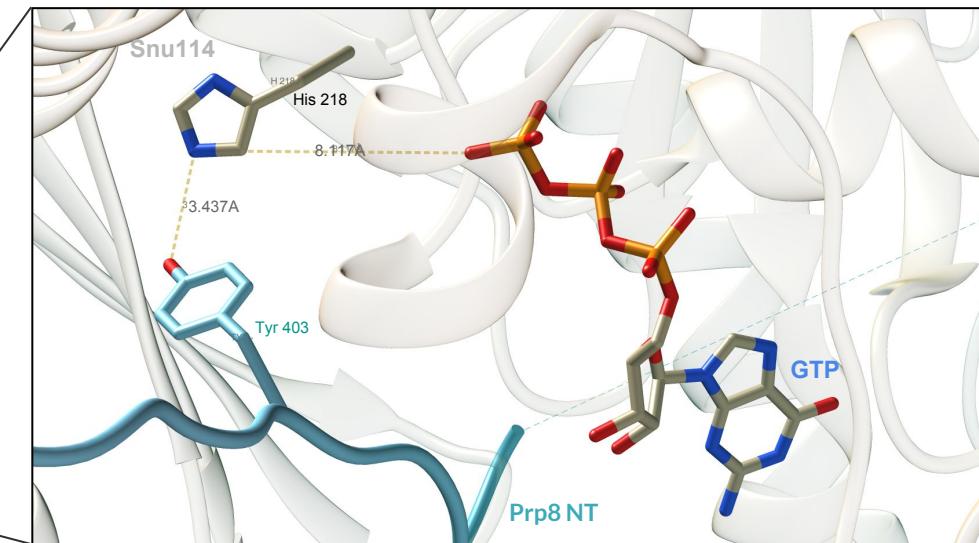
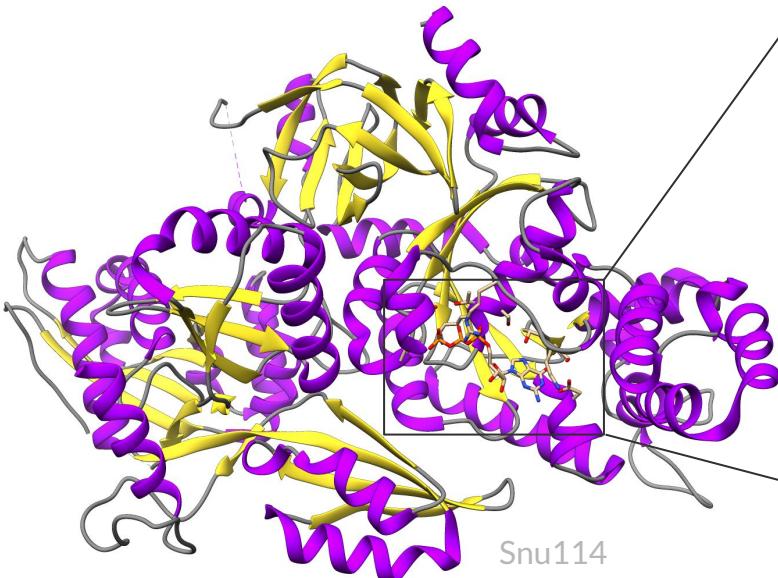
N-terminal



PDB code: 5GAM
Resolution: 3.7A

Snu114

In most GTPases the glutamine residue hydrolyses the phosphate ester...



...but in Snu114 the glutamine is replaced by histidine, which is hydrogen bonded to tyrosine.

N-terminal

Tyr 403

	361	420
H.sapiens	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
C.lupus	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
M.musculus	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
M.mulatta	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
G.gallus	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
X.tropicalis	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
D.rerio	VRDINLQDED WNEFNDINKI	IIRQPIRTEY KIAFPYLYNN LP..HHVHLT WYHTPNVVFI
D.melanogaster	IKDHNVGDED WNEFNDINKV	IIRQPIRTEY RIAFPYLYNN MP..HFVHLS WYHTPNVVYI
C.elegans	VKDLHT.DED WNEFNDINKV	IIRAPIRTEY RIAFPFMYNN LISSLPVQVS WYHTPSVVFI
A.thaliana	HRDMEKGDED WNEFNDINKL	IIRSPLRTEY KVAFPHLYNN RP..RKVKLC VYHTPMVMYI
S.pombe	YKDEAPEMED WNEFNDIYKL	IIRHPIKTEY RIAFPYLYNS RA..RSVALS EYHQPSNVFV
N.crassa	YKDIDPNDED FGEFNAMDRI	IFRNPIRTEC RVAYPHLYNA LP..RSVQLS VHSYQPQVVYT
S.cerevisiae	YP..REEEEED YNEFNSIDRV	IFRVPIRSEY KVAFPHLYNS RP..RSVRIP WYNNPVCII

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

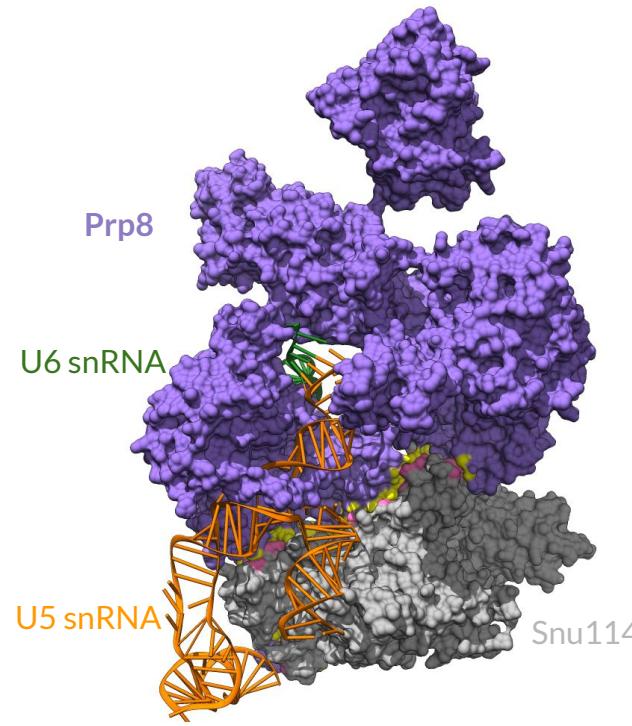
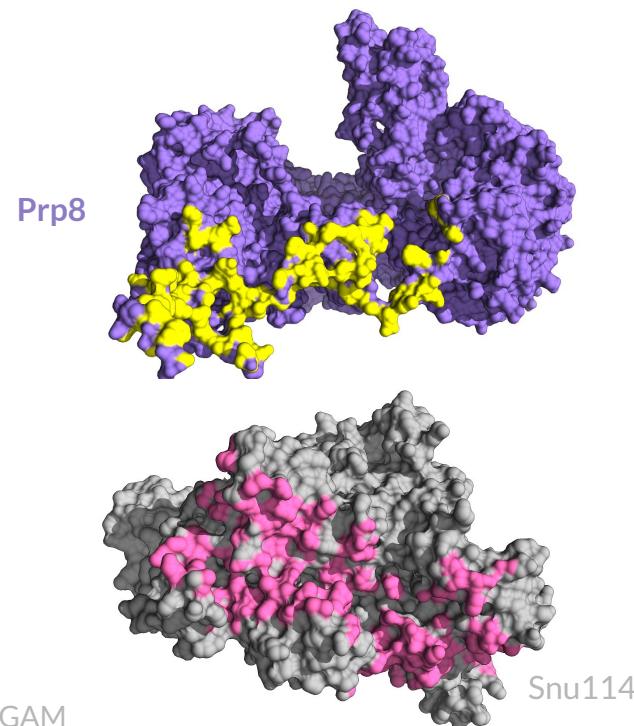
Linker

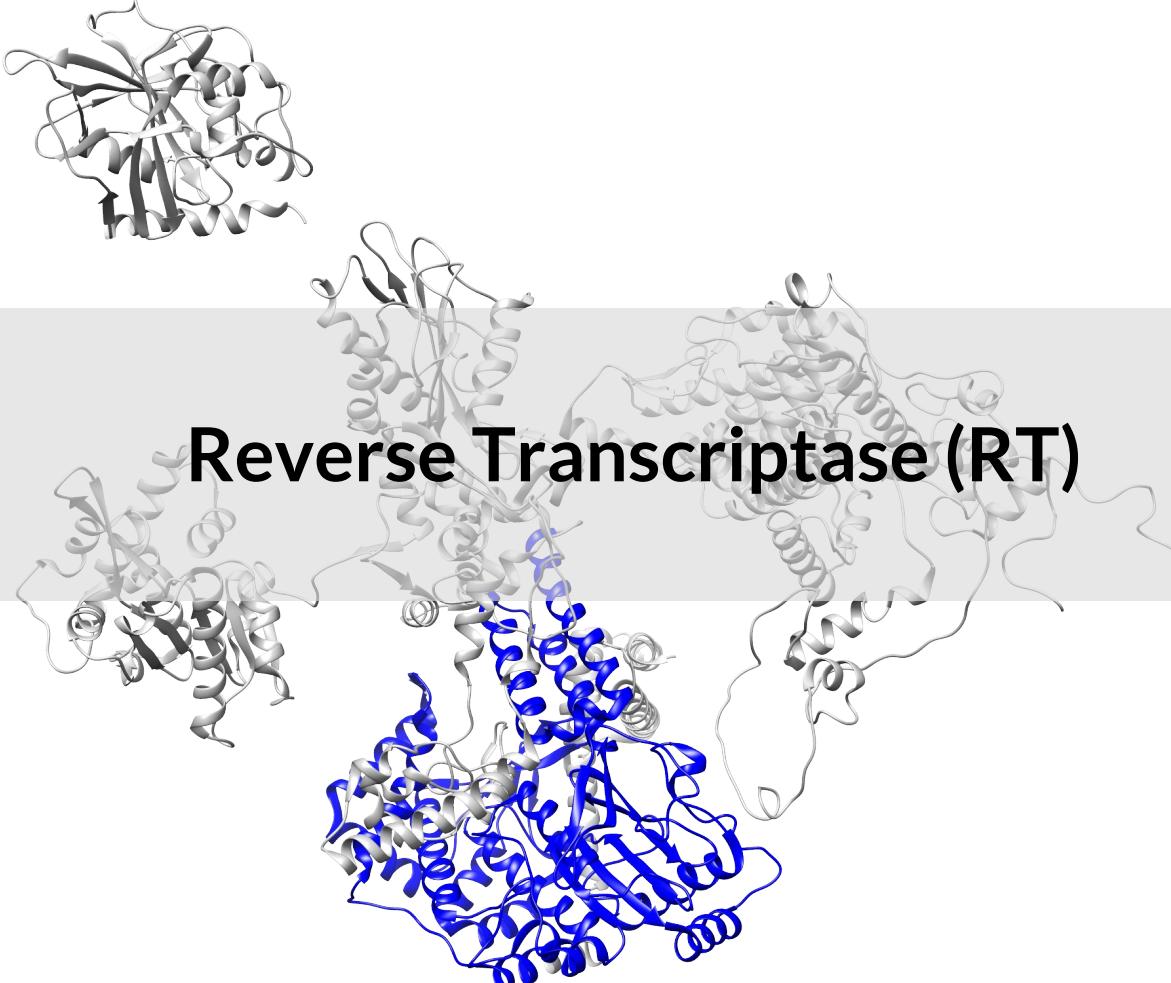
Endonuclease

RNase H-like

Jab1/MPN

N-terminal - Snu114





A ribbon diagram of the Reverse Transcriptase (RT) structure. The diagram is composed of grey and blue ribbon segments, representing different domains or regions of the protein. A central, more compact blue region is surrounded by a larger, more extended grey region. The entire structure is set against a white background.

Reverse Transcriptase (RT)

N-domain

RT fingers/palm

Thumb/X

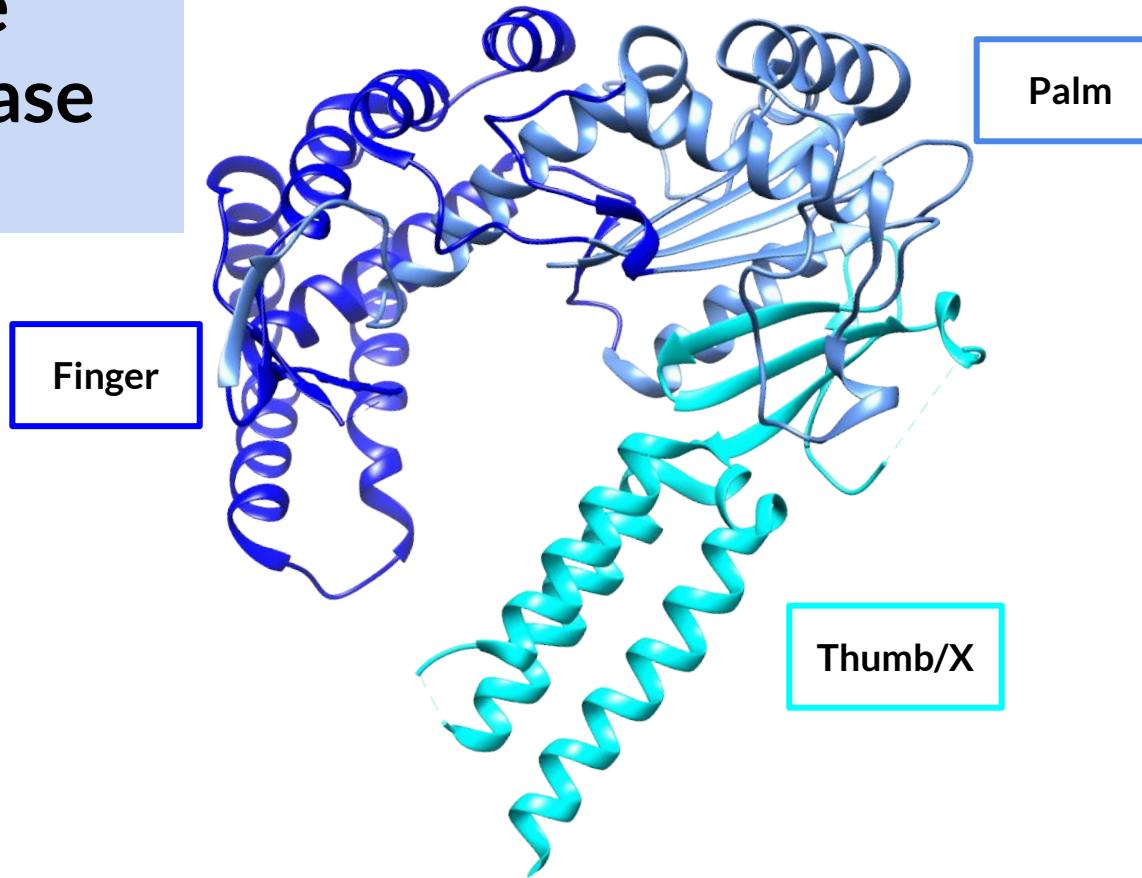
Linker

Endonuclease

RNase H-like

Jab1/MPN

Reverse transcriptase (RT)



PDB code: 4I43
Resolution: 2A

N-domain

RT fingers/palm

Thumb/X

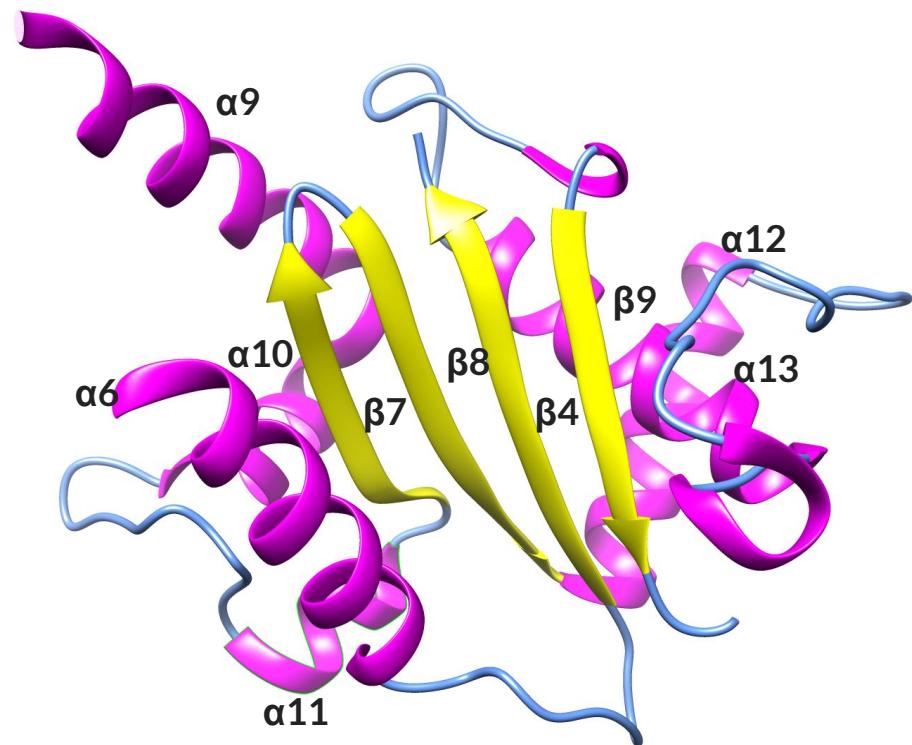
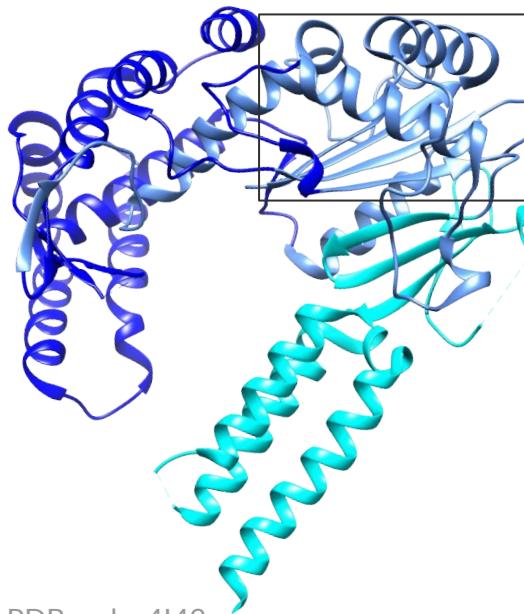
Linker

Endonuclease

RNase H-like

Jab1/MPN

RT palm



PDB code: 4I43
Resolution: 2A

N-domain

RT fingers/palm

Thumb/X

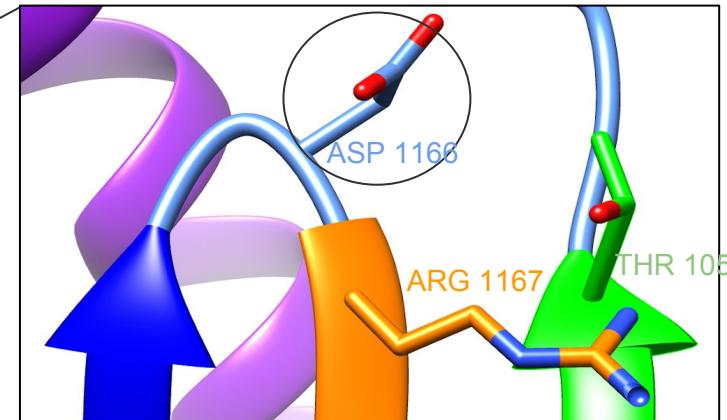
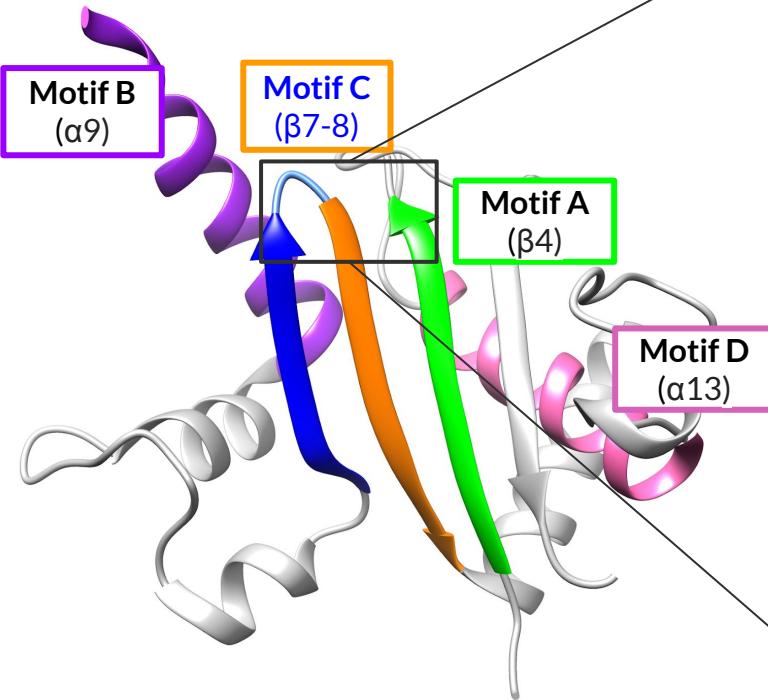
Linker

Endonuclease

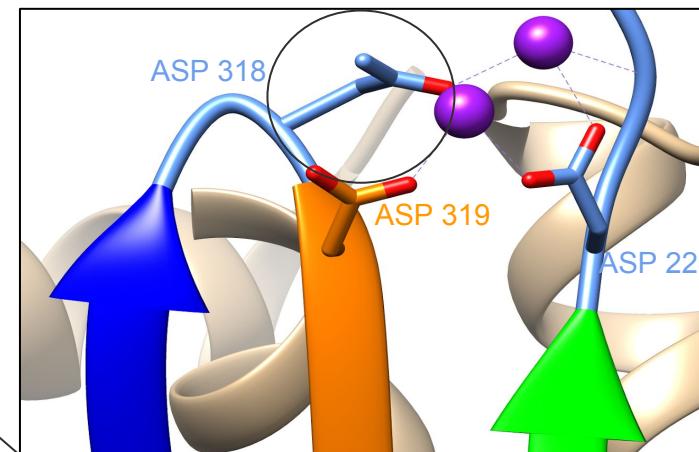
RNase H-like

Jab1/MPN

RT palm



Prp8
PDB code: 4I43



HCV
PDB code: 1NB6

RT palm

Thr 1053

Motif A

	1021							1080
H. sapiens	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETSEGECSV	MLESRFEKMY		
C. lupus	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETSEGECSV	MLESRFEKMY		
M. musculus	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETSEGECSV	MLESRFEKMY		
M. mulatta	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETSEGECSV	MLESRFEKMY		
G. gallus	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETSEGECSV	MLESRFEKMY		
X. tropicalis	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WDTTEGECSV	MLESRFEKMY		
D. rerio	LWYEADKRRRL	FPPWIKPADT	EPPPLLVYKW	CQGINNLQDV	WETAEGECNV	MLESRYEKMY		
D. melanogaster	LWYEADKRRRL	FPPWIKPSDT	EPPPLLAYKW	CQGINNLQDV	WDVGEGECSV	LLESRFEKLY		
C. elegans	LWYEADKRRRL	FPAWVKPGDT	EPPPLLTYKW	CQGLNNLQDV	WETSEGECSV	IMETKLEKIA		
A. thaliana	LWYEGDKRHL	FPNWIKPADS	EPPPLLVYKW	CQGINNLQGI	WDTSDGQCVV	MLQTKEFKLF		
S. pombe	LWFEADRRHL	FPSWVKPSDS	EPPPLLVYKW	CQGINNLTDV	WETSNGECNV	LMETRLSKVF		
N. crassa	LWYQADQRHL	FPAWIKPSDS	EVPPPLLVYKW	AQGINNLDRV	WETANGECNV	MIETQLSKVY		
S. cerevisiae	LWYEADQRKL	FPNWIKPSDS	EIPPLLVYKW	TQGINNLSEI	WDVSRGQSAV	LLETTLGEMA		

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

Jab1/MPN

RT palm

	1141	Asp 1166		Arg 1167		Motif C	1200
H. sapiens	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DIATEAAHPI	RLFCRYIDRI	HIFFRFTADE	
C. lupus	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DIATEAAHPI	RLFCRYIDRI	HIFFRFTADE	
M. musculus	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DIATEAAHPI	RLFCRYIDRI	HIFFRFTADE	
M. mulatta	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DIATEAAHPI	RLFCRYIDRI	HIFFRFTADE	
G. gallus	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DIATEVAHPI	RLFCRYIDRI	HIFFRFTADE	
X. tropicalis	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DVATESAHPI	RLFCRYIDRI	HIFFRFSADE	
D. rerio	LVMDLLVLGL	HRASEMAGPP	QMPNDLFSFQ	DTATESAHPI	RLYCRYIDRI	HIFFRFSADE	
D. melanogaster	LVLDLLVLGL	HRSSEMAGPP	QMPNDLTFQ	DTVTETAHPI	RLYCRYVDRI	HLFFRFSAAE	
C. elegans	LVLDLLVLGL	RRASEIAGPP	QCPNEFLQFQ	DVATEIGHPI	RLYCRYIDRV	WIMFRFSADE	
A. thaliana	LVLDLLLLGL	TRASEIAGPP	QRPNEFMTYW	DTKVETRHPI	RLYSRYIDKV	HIMFKFTHEE	
S. pombe	LVLDLLILGL	QRATEIAGPA	DAPNDLHFQ	DQATETSHPI	RLYTRYIDKV	YIMFRFTDEE	
N. crassa	LILDLPLLGP	QRASEIAGPP	HAPNDLQFK	DRETETRHPI	RLYTRYIDKI	WVFLRFTADE	
S. cerevisiae	LVIDLLLLGQ	ERATDLAGPA	NNPNEFMQFK	SKEVEKAHPI	RLYTRYLDRI	YMLFHFEDE	

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

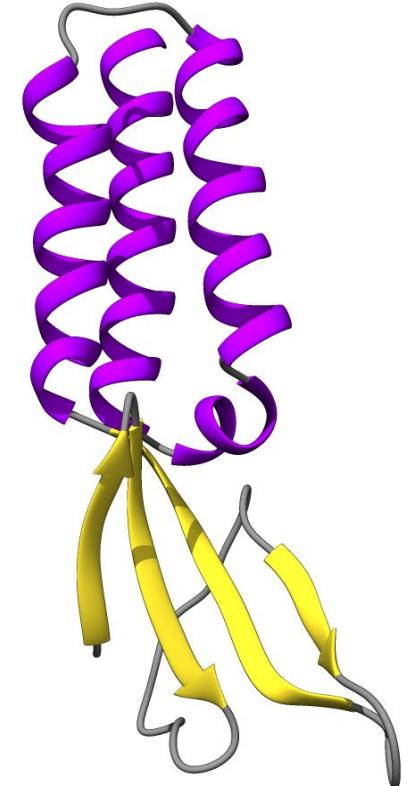
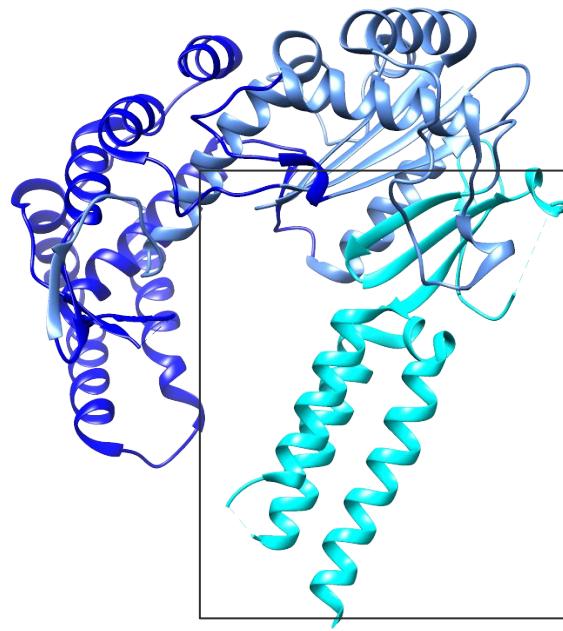
Linker

Endonuclease

RNase H-like

Jab1/MPN

Thumb/X



PDB code: 4I43
Resolution: 2A

N-domain

RT fingers/palm

Thumb/X

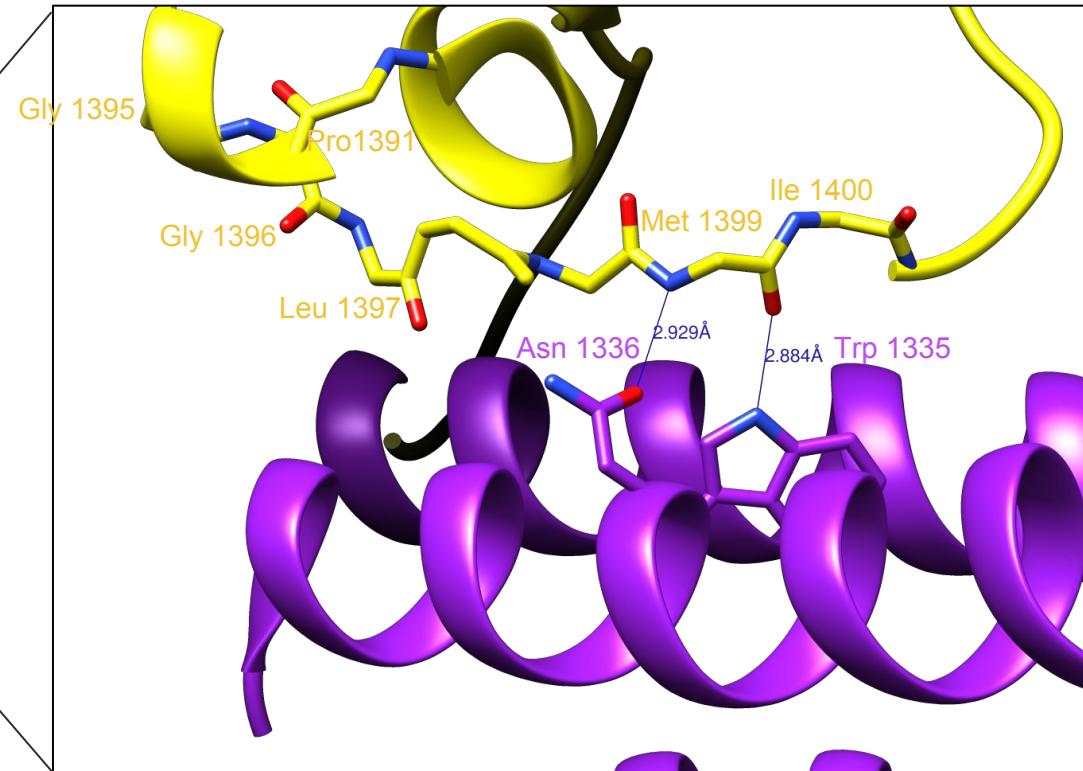
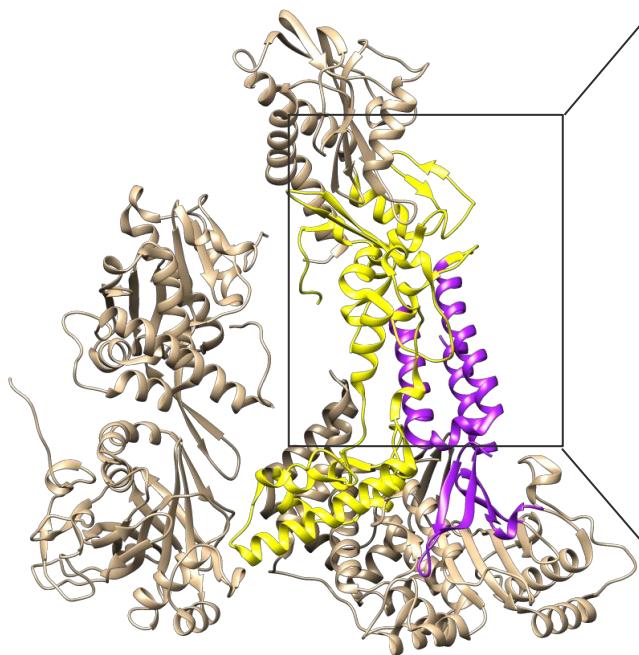
Linker

Endonuclease

RNase H-like

Jab1/MPN

Thumb/X - Linker



PDB code: 4I43
Resolution: 2A

■ Linker ■ Thumb/X

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

Jab1/MPN

Thumb/X - Linker

Trp 1335

Asn 1336

Thumb region

	1321						1380
H. sapiens	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
M. mulatta	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
C. lupus	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
M. musculus	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
G. gallus	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
D. rerio	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
D. melanogaster	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
C. elegans	MSSGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	
S. cerevisiae	MASGSTTFTK	VAAK W NTSLI	SLFTYFREAI	VATEPLLDIL	VKGTRRIQNR	VKLGLNSKMP	
S. pombe	MSSGSTTFTK	IANK W NTALI	ALMTYYREAA	ISTPELLDLL	VKCESKIQTR	VKISLNSKMP	
N. crassa	MSSGSTTFTK	IANK W NTALI	ALFTYYREAA	VSTVNLLDTI	VKCETKIQTR	VKIGLNSKMP	
A. thaliana	MSSGSTTFTK	IVNK W NTALI	GLMTYFREAT	VHTQELLDLL	VKCENKIQTR	VKIGLNSKMP	
X. tropicalis	MASGSTTFTK	IVNK W NTALI	GLMTYFREAV	VNTQELLDLL	VKCENKIQTR	IKIGLNSKMP	

Negatively charged

Non-polar

Positively charged

Polar

Sequence alignment based on PFAM domains

Thumb/X - Linker

Gly 1395
Gly 1396
Leu 1397

Linker region

Sequence alignment based
on PFAM domains

	1381	SRFPPVVVFYT	PKELGGLGML	SMGHVLIPQSQ	DLRWSkqtdv	githfrsgms	heedqlipnl	1440
H. sapiens								
M. mulatta								
C. lupus								
M. musculus								
G. gallus								
D. rerio								
D. melanogaster								
C. elegans								
S. cerevisiae		TRFPPAVFYT	PKELGGLGMI	SASHILIPAS	DLSWSkqtdt	githfragmt	hedeqlipti	
S. pombe		SRFPPAVFYS	PKELGGLGML	SMGHVLIPQSQ	DLRWSkqtdt	githfrsgmt	tngehlipnl	
N. crassa		SRFPPAVFYT	PKELGGLGMI	SGSHILIPTS	DKRWSkqtdl	gvthyragms	hdeetlipni	
A. thaliana		SRFPPVIFYT	PKEIGGLGML	SMGHILIPQSQ	DLRYSnqtdv	gvshfrsgms	heedqlipnl	
X. tropicalis		SRFPPVVVFYT	PKELGGLGML	SMGHVLIPQSQ	DLRWSkqtdv	githfrsgms	heedqlipnl	

Pro 1384
Pro 1385
Pro 1391

Gly 1398
Met 1399
Leu/Ile 1400

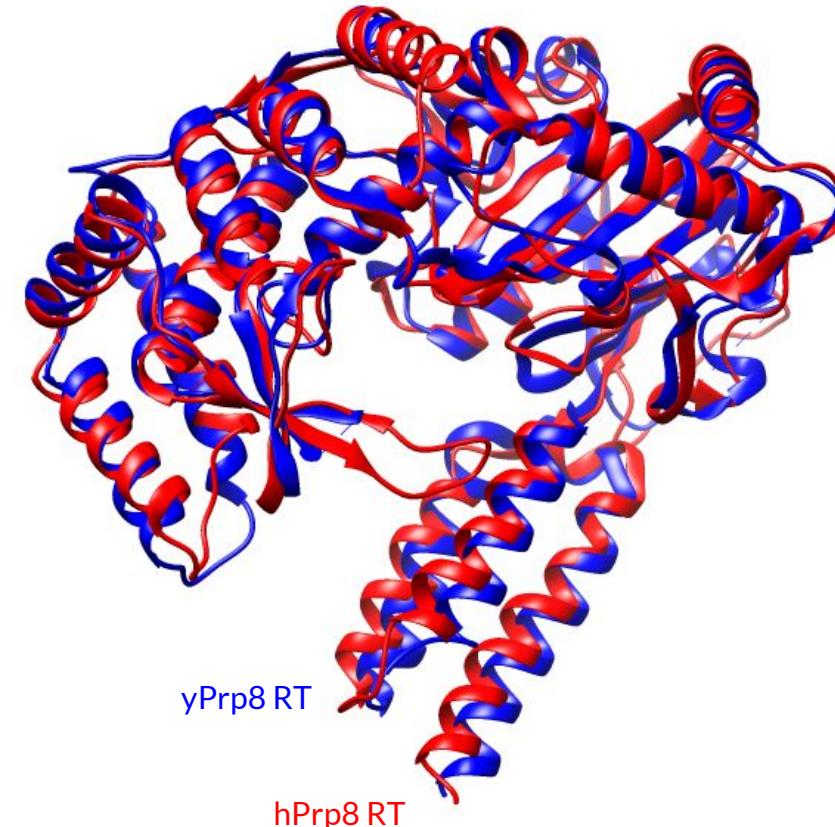
Negatively charged Non-polar
Positively charged Polar

Reverse transcriptase (RT)

STAMP VALUES

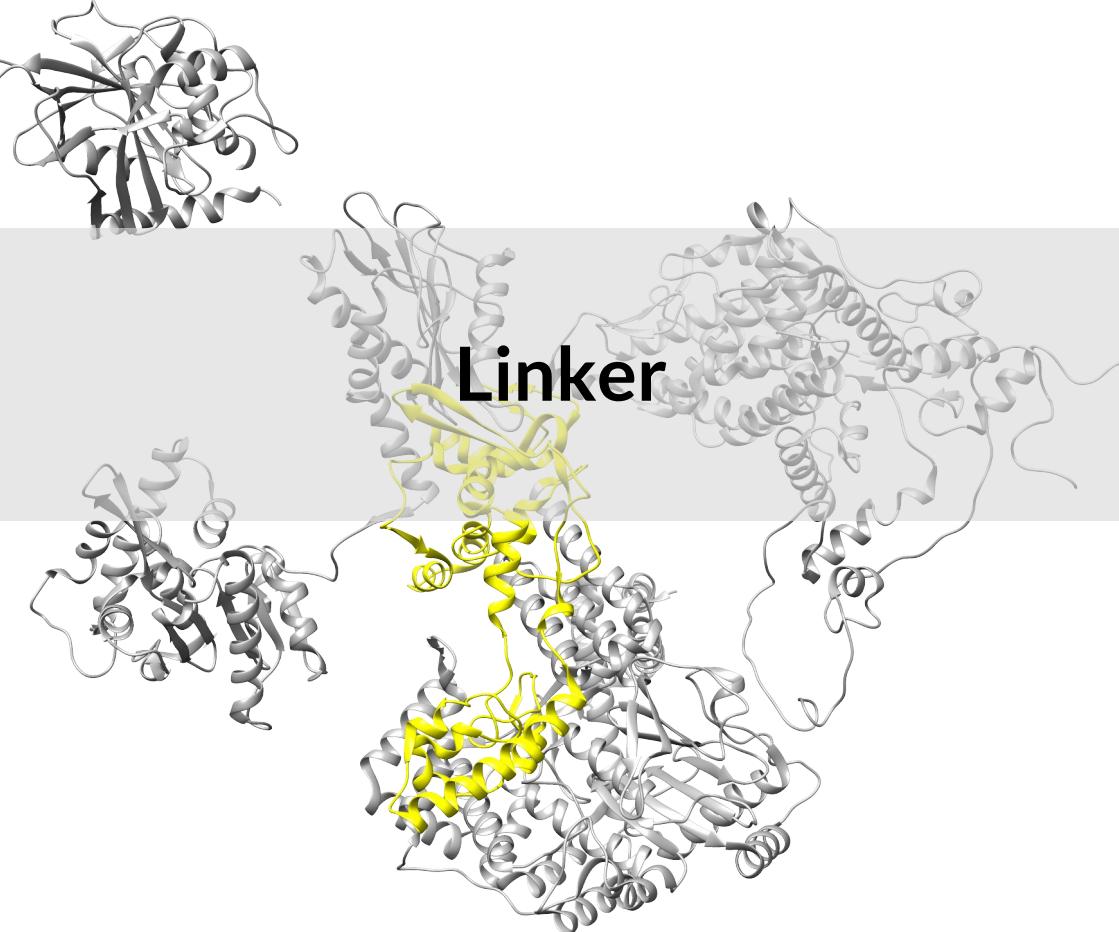
Sc 7.21

RMS 1.65



PDB code: 5MQF
Resolution: 5.9A

PDB code: 4I43
Resolution: 2A



Linker

N-domain

RT fingers/palm

Thumb/X

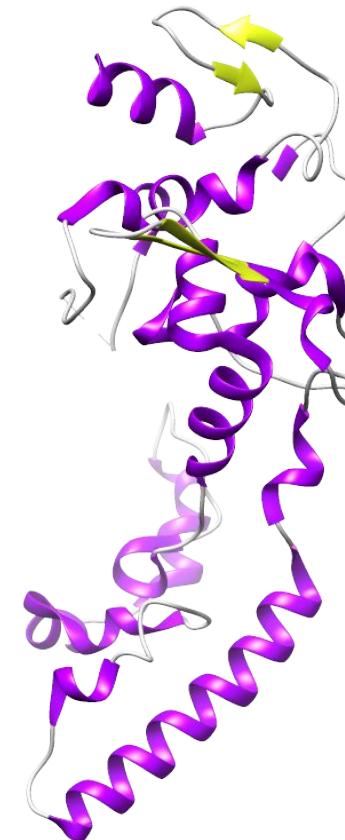
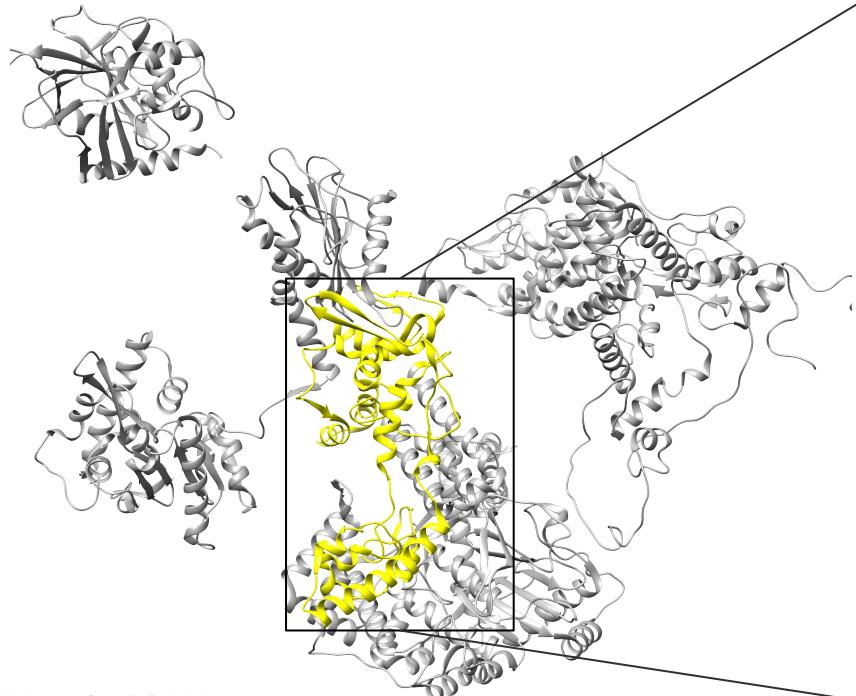
Linker

Endonuclease

RNase H-like

Jab1/MPN

Linker



PDB code: 5GAM
Resolution: 3.7Å

N-domain

RT fingers/palm

Thumb/X

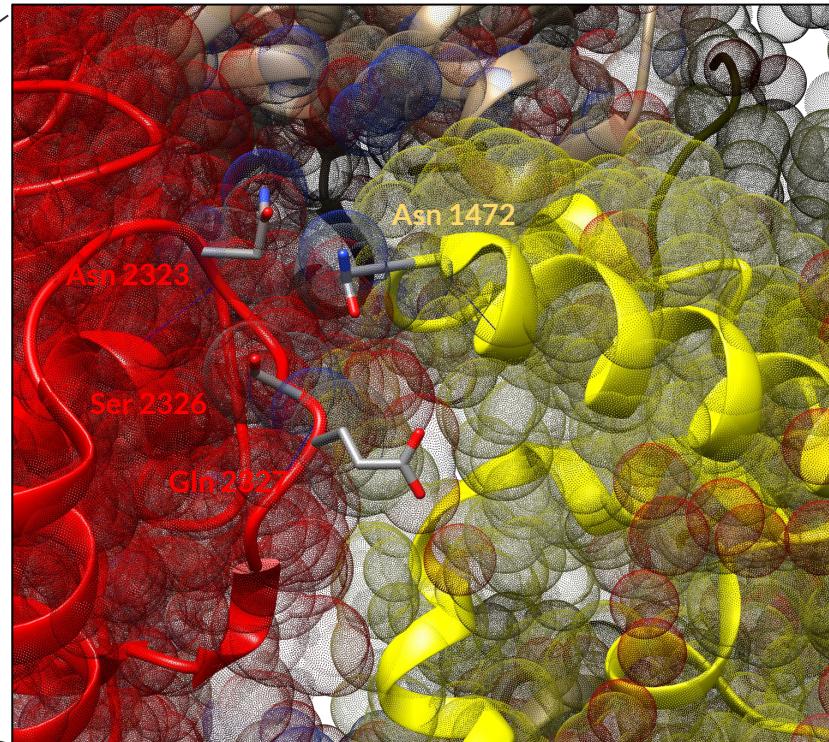
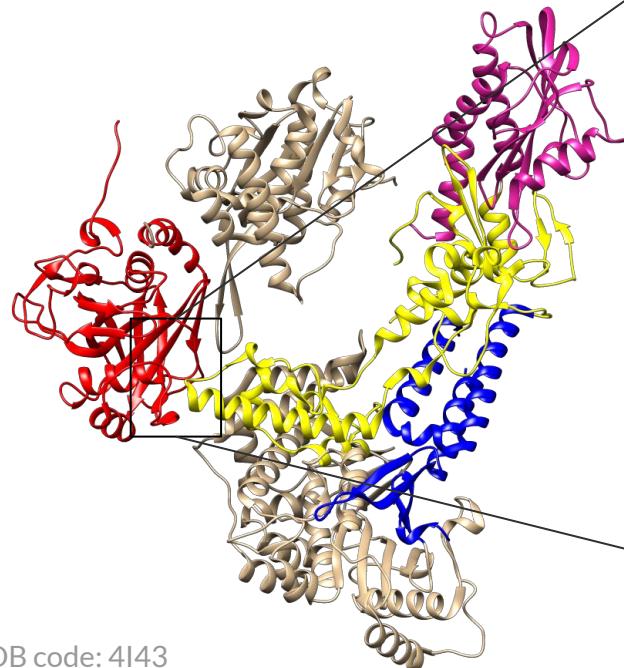
Linker

Endonuclease

RNase H-like

Jab1/MPN

Linker- Jab1/MPN



PDB code: 4I43
Resolution: 2A

Linker

Thumb/X

Jab1/MPN

Endonuclease

Linker- Jab1/MPN

Jab1/MPN region

	2281											2340	
H. sapiens	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
M. mulatta	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
C. lupus	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
M. musculus	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
G. gallus	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
D. rerio	pswdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
D. melanogaster	snwdgektIV	ITCSFTPGSC	SLTAYKLTPS	GFEWGSKNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML
C. elegans	iswdgektV	ITCSFTPGSV	SLTAYKLTPS	GYEWGKANTD	.	K	G	N	N	P	K	GYM	PTHYEKVQML
S. cerevisiae	fadkkrdCID	ISIFSTPGSV	SLSAYNLTDE	GYQWGEENKD	I	M	N	V	L	S	E	GFE	PTFSTHAQLL
S. pombe	pewdtka.VT	LTVSYIPGSI	SLAAAYTVSKE	GIEWGSKNMD	I	N	S	D	E	A	I	GYE	PSMAEKCQLL
N. crassa	nwdkqnt.LT	VAVSFTPGSV	SLSAWALTPQ	GFKWGVENKD	I	A	S	D	Q	P	Q	GFT	TSMGEKRQLL
A. thaliana	kqwdaekCII	LTCSFTPGSC	SLTSYKLTQA	GYEWGRLNKD	.	T	G	S	N	P	H	GYL	PTHYEKVQML
X. tropicalis	pawdgektII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.	K	G	N	N	P	K	GYL	PSHYERVQML

Negatively charged

Non-polar

Positively charged

Polar

Asn 2323
Ser 2326

Glu 2327

Sequence alignment based on PFAM domains

Linker- Jab1/MPN

Linker region

Asn 1472

	1441	1500
H. sapiens	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
C. lupus	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
M. musculus	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
M. mulatta	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
G. gallus	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
X. tropicalis	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
D. rerio	SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAIQ N R	
D. melanogaster	SKQTDVG.IT HFRSGMSHDE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEANAQ N R	
C. elegans	MQQTEAGGVT HFRSGMSHDE DQLIPNLYRY IQPWEAEFVD SVRVWALEYAL KRQEANAQ N R	
A. thaliana	SNQTDVG.VS HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWALEYAL KRQEAQ N R	
S. pombe	SKQTDTG.IT HFRSGMTTNG EHLIPNLYRY IQPWESEFID SQRVWALEYAM KRQEALQQ N R	
N. crassa	SKQTDLG.VT HYRAGMSHDE ETLIPNIFRY IIPWEAEFID SQRVWTEYSQ KRLEANQQ N R	
S. cerevisiae	SKQTDTG.IT HFRAGMTHED EKLIPTIFRY ITTWEAEFLD SQRVWALEYAT KRQEAIQQ N R	

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

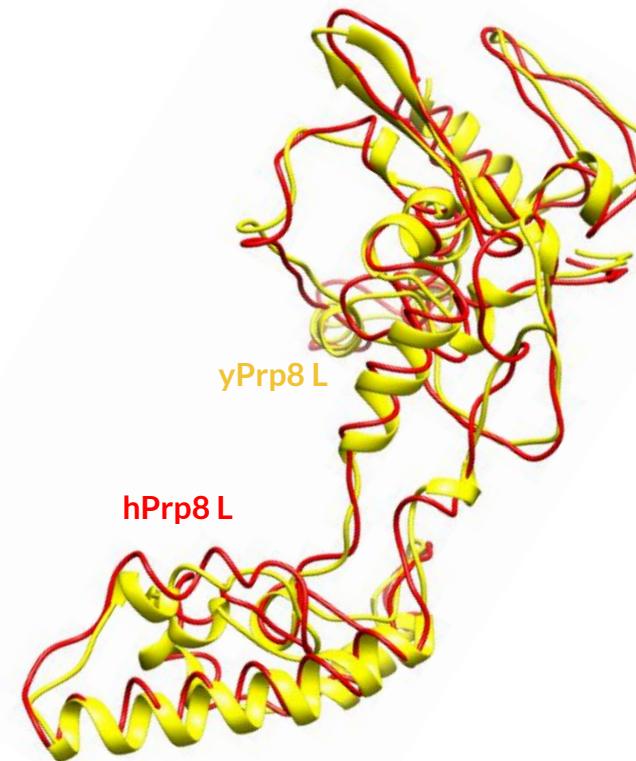
Jab1/MPN

Linker

STAMP VALUES

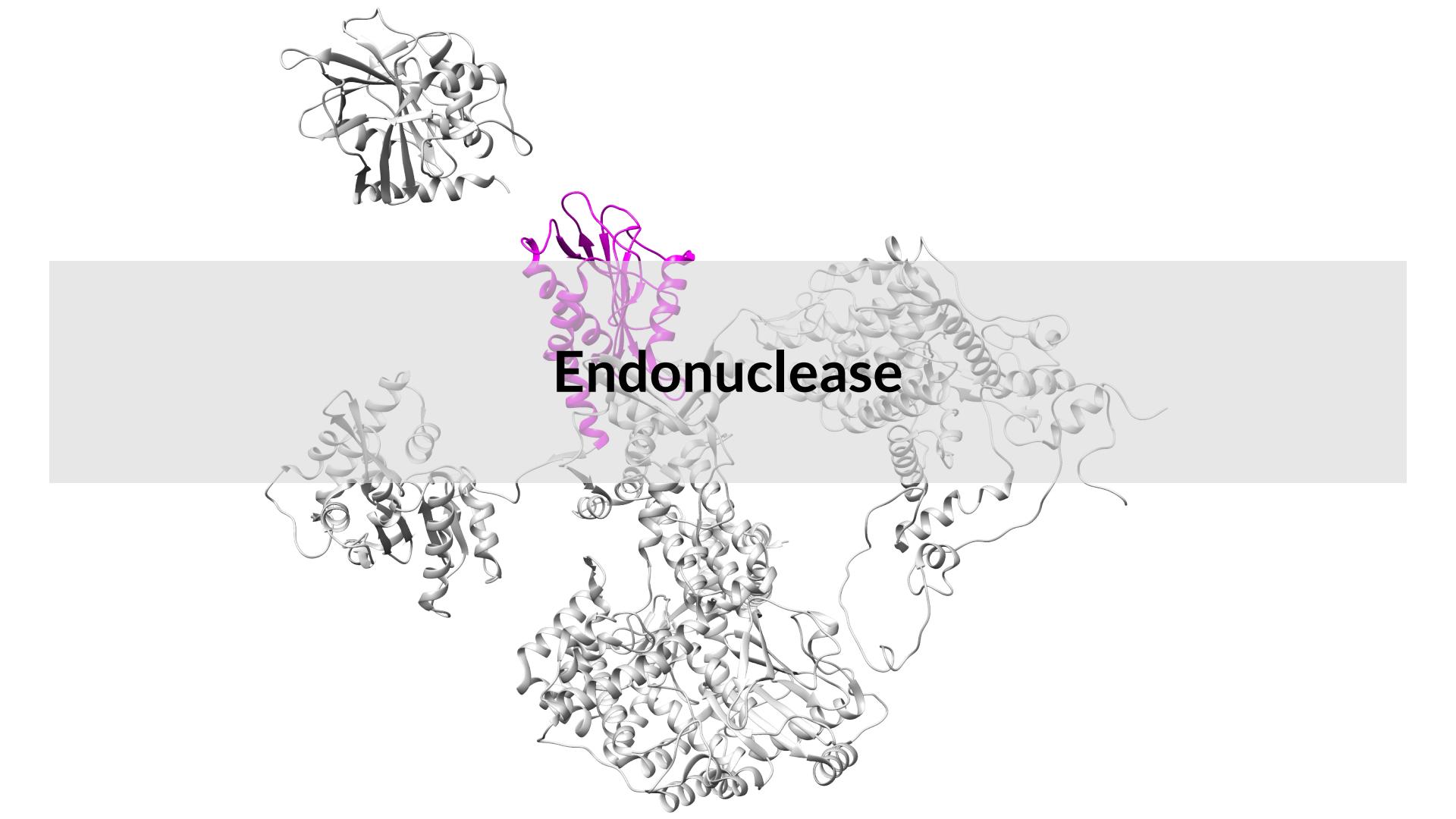
Sc 8.43

RMS 1.42



PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A



Endonuclease

N-domain

RT fingers/palm

Thumb/X

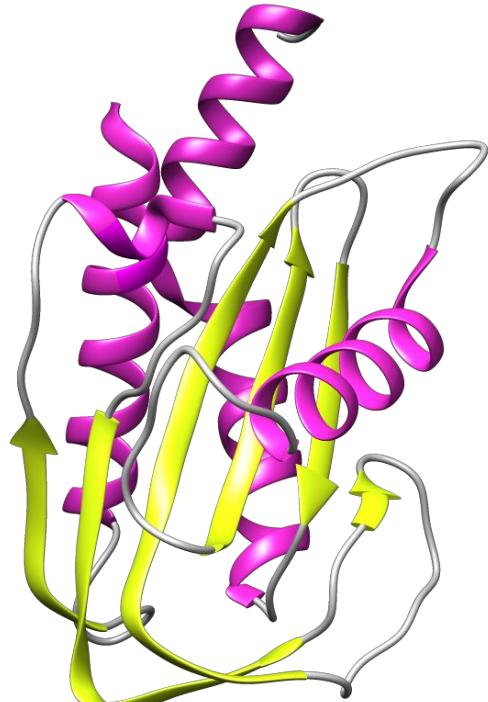
Linker

Endonuclease

RNase H-like

Jab1/MPN

Endonuclease



PDB code: 4I43
Resolution: 2A

N-domain

RT fingers/palm

Thumb/X

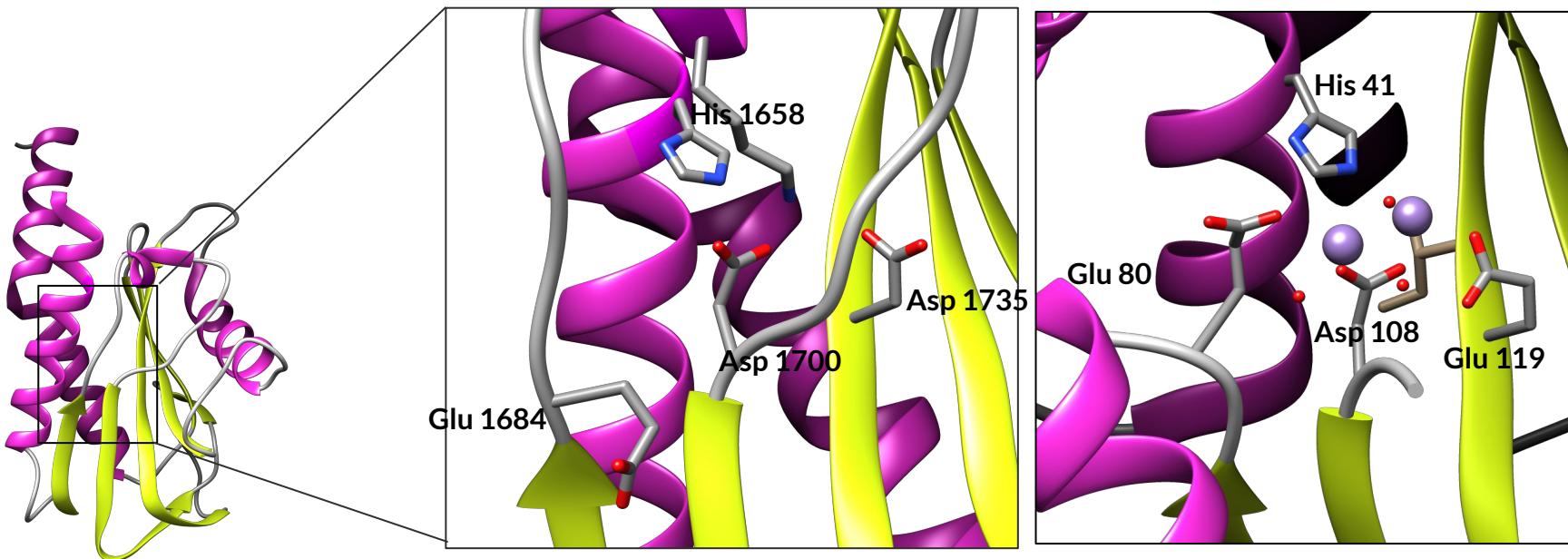
Linker

Endonuclease

RNase H-like

Jab1/MPN

Endonuclease



Prp8
PDB code: 4I43

Influenza Virus
PDB code: 2W69

N-domain

RT fingers/palm

Thumb/X

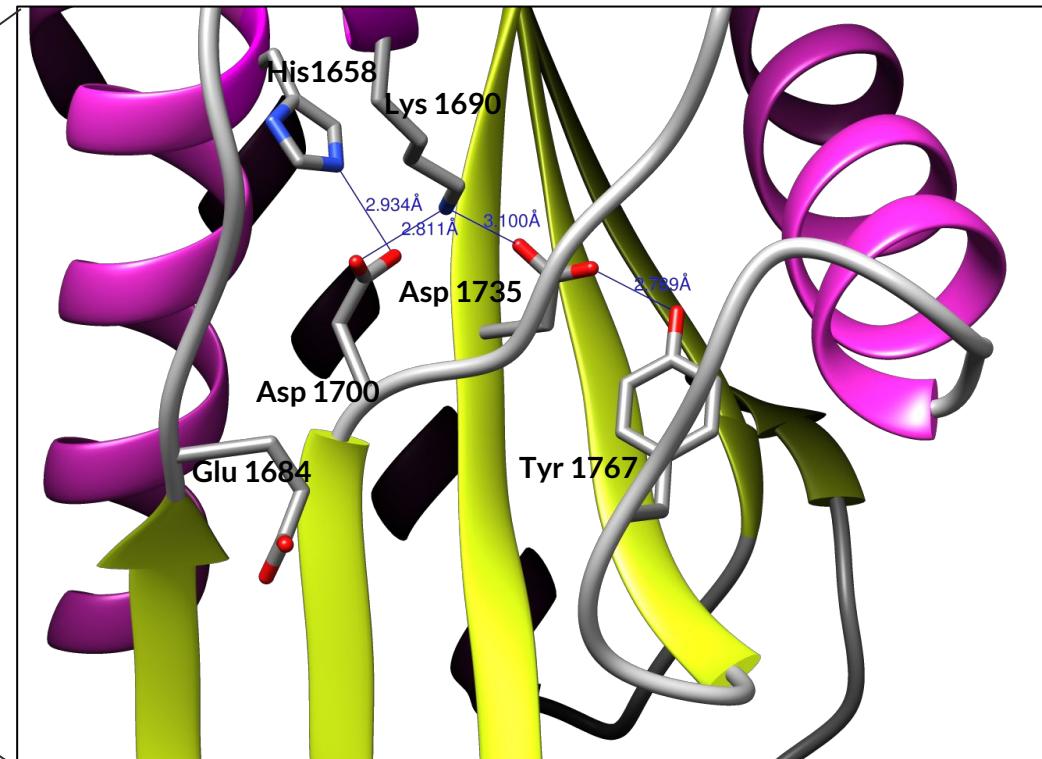
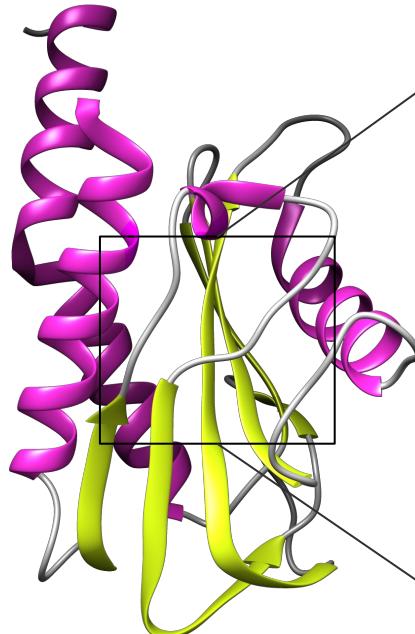
Linker

Endonuclease

RNase H-like

Jab1/MPN

Endonuclease



PDB code: 4I43
Resolution: 2Å

Endonuclease

His 1658

	1621						1680
H. sapiens	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
M. mulatta	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
C. lupus	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
M. musculus	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
G. gallus	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
D. rerio	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
D. melanogaster	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	
C. elegans	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	VVMDLCQVFD	QELdaleiqt	
S. cerevisiae	VGFLVQLDLT	GIFLHGKIPT	LKISLIQIFR	AHLWQKI HES	IVFDICQILD	GELdvlqies	
S. pombe	VGFQVQLDLT	GIMMHGKIPT	LKISLIQIFR	SHLWQKI HES	VVWDLCQVLD	QELeslqiet	
N. crassa	VGFQVQLDLT	GIFLHGKIPT	LKISLIQIFR	AHLWQKI HES	VVMDLCQVFD	QELealsies	
A. thaliana	VGFQVQLDLT	GIYMHGKIPT	LKISLIQIFR	AHLWQKI HES	VVMDLCQVLD	QELepleiet	
X. tropicalis	VGFQVQLDLT	GIFMHGKIPT	LKISLIQIFR	AHLWQKI HES	IVMDLCQVFD	QELdaleiet	

Negatively charged

Non-polar

Positively charged

Polar

Sequence alignment based on PFAM domains

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

Jab1/MPN

Endonuclease

Glu 1684

Lys 1690

Asp 1700

	1681		1740
H. sapiens	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
C. lupus	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
M. musculus	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
M. mulatta	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
G. gallus	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
X. tropicalis	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
D. rerio	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FASYKWNVSR	
D. melanogaster	WQKIHESIVM DLCQVFDQEL DALEIETVQK	ETIHP R KSYK MNSSCAD I LL FPAYKWNVSR	
C. elegans	WQKIHESVVM DLCQVFDQEL DALEIQT V QK	ETIHP R KSYK MNSSCAD V LL FAQYKWNVSR	
A. thaliana	WQKIHESVVM DLCQVLDQEL EPLEIETVQK	ETIHP R KSYK MNSSCAD V LL FAAHKWPMSK	
S. pombe	WQKIHESVWW DLCQVLDQEL ESLQIETVQK	ETIHP R KSYK MNSSCAD I LL LAAYKWNVSR	
N. crassa	WQKIHESVVM DLCQVFDQEL EALSIESVQK	ETIHP R KSYK MNSSCAD I QL FASHKWNVTR	
S. cerevisiae	WQKIHESIVF DICQILDGEL DVLQIESVTK	ETVHPR K SYK MNSSAAD I TM ESVHEWEVSK	

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

Endonuclease

Asp 1735

Tyr 1767

1741

1800

H. sapiens	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
C. lupus	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
M. musculus	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
M. mulatta	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
G. gallus	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
X. tropicalis	PSLLADSKDV	MDSTTTQKYW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
D. rerio	PSLLADSKDV	MDSTTTQFW	IDIQLRWGDY	DSHDIERYAR	AKFLDYTTDN	MSIYPSPTGV
D. melanogaster	PSLLADTKDT	MDNTTTQKYW	LDIQLRWGDY	DSHDVERYAR	AKFLDYTTDN	MSIYPSPTGV
C. elegans	PSLMADSKDV	MDNTTTQKYW	LDVQLRWGDY	DSHDVERYAR	AKFLDYTTDN	MSIYPSPTGV
A. thaliana	PSLIAESKDV	FDQKASNKYW	IDVQLRWGDY	DSHDIERYTK	AKFMDYTTDN	MSIYPSPTGV
S. pombe	PSLLNDNRDV	LDNTTNKYW	IDVQLRFGDY	DSHDIERYTR	AKFLDYSTDA	QSMYPSPTGV
N. crassa	PSLLFDTKDV	IESTTTNKFW	IDVQLRYGDY	DSHDIERYVR	AKYLDYTTDS	MSLYPSPTGL
S. cerevisiae	PSLLHETNDS	FKGLITNKMW	FDVQLRYGDY	DSHDISRYVR	AKFLDYTTDN	VSMYPSPTGV

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

Endonuclease

	1681		1740
H.sapiens	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
C.lupus	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
M.musculus	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
M.mulatta	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
G.gallus	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
X.tropicalis	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
D.rerio	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FASYKWNVSR
D.melanogaster	WQKIHESIVM	DLCQVFDQEL DALEIETVQK	E TIHPRKSYK MNSSCADILL FPAYKWNVSR
C.elegans	WQKIHESVVM	DLCQVFDQEL DALEIQTVQK	E TIHPRKSYK MNSSCADVLL FAQYKWNVSR
A.thaliana	WQKIHESVVM	DLCQVLDQEL EPLEIETVQK	E TIHPRKSYK MNSSCADVLL FAAHKWPMSK
S.pombe	WQKIHESVWW	DLCQVLDQEL ESLQIETVQK	E TIHPRKSYK MNSSCADILL LAAYKWNVSR
N.crassa	WQKIHESVVM	DLCQVFDQEL EALSIESVQK	E TIHPRKSYK MNSSCADIQL FASHKWNVTR
S.cerevisiae	WQKIHESIVF	DICQILDGEL DVLQIESVTK	E TVHPRKSYK MNSSAADITM ESVHEWEVSK
	1741		1800
H.sapiens	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
C.lupus	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
M.musculus	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
M.mulatta	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
G.gallus	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
X.tropicalis	PSLLADSKDV	MDSTTTQKYW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
D.rerio	PSLLADSKDV	MDSTTTQKFW IDIQLRWGDY	D SHDIERYAR AKFLDYTTDN MSIYPSPTGV
D.melanogaster	PSLLADTKDT	MDNTTTQKYW LDIQLRWGDY	D SHDVERYAR AKFLDYTTDN MSIYPSPTGV
C.elegans	PSLMADSKDV	MDNTTTQKYW LDVQLRWGDY	D SHDVERYAR AKFLDYTTDN MSIYPSPTGV
A.thaliana	PSLIAESKDV	FDQKASNKYW IDVQLRWGDY	D SHDIERYTK AKFMDYTTDN MSIYPSPTGV
S.pombe	PSLLNDNRDV	LDNTTTNKYW IDVQLRFGDY	D SHDIERYTR AKFLDYSTDQ QSMYPSPTGV
N.crassa	PSLLFDTKDV	IESTTTNKFW IDVQLRYGDY	D SHDIERYVR AKYLDYTTDS MSLYPSPTGL
S.cerevisiae	PSLLHETNDS	FKGLITNKMW FDVQLRYGDY	D SHDISRYVR AKFLDYTTDN VSMYPSPTGV

Loops sequence highly conserved:

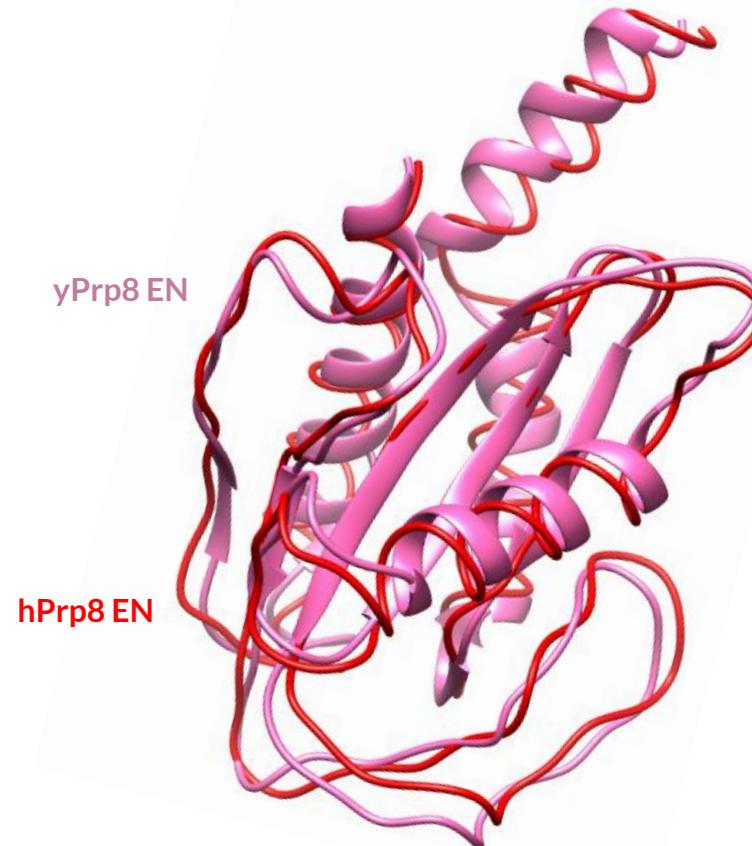
- Residues 1685-1699
- Residues 1762-1771

Clustal Sequence alignment

Endonuclease

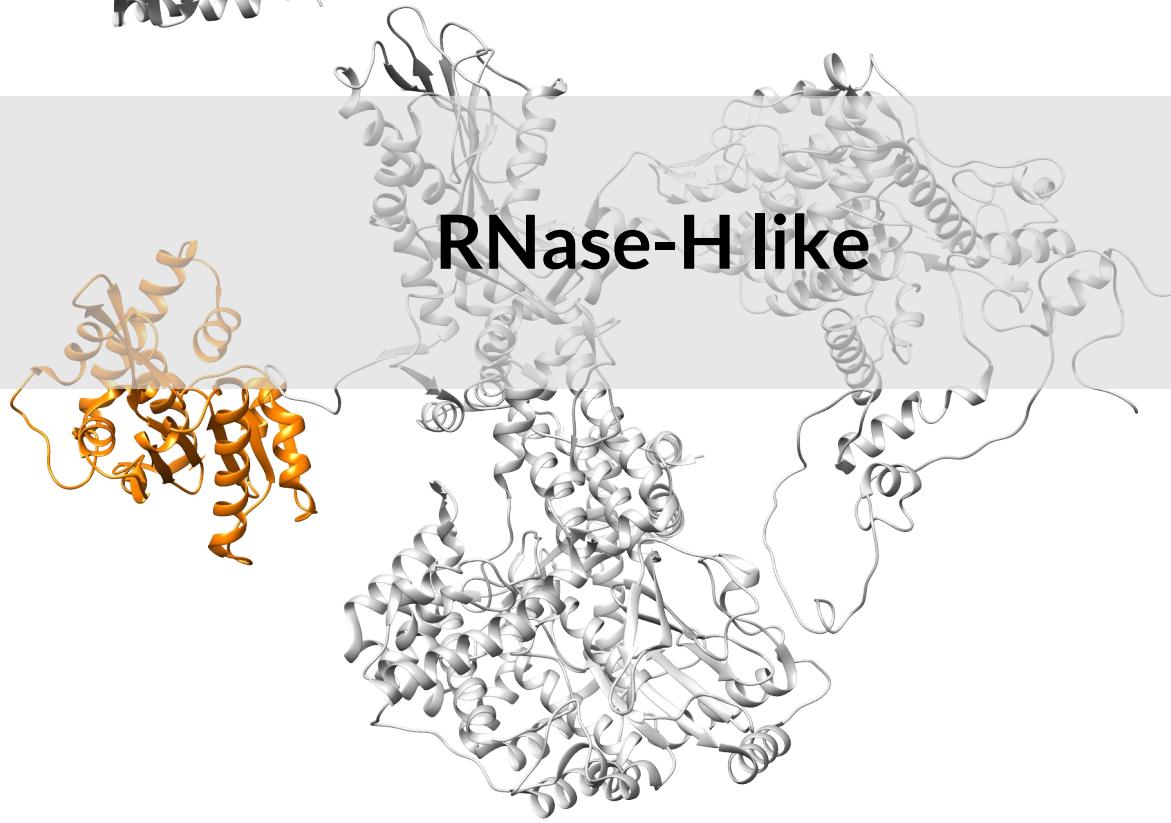
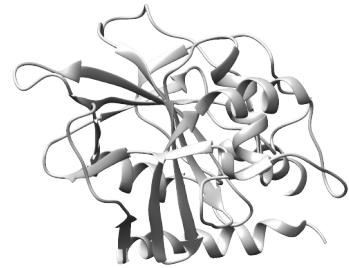
STAMP VALUES

- Sc 9.12
- RMS 1.19



PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A



RNase-H like

N-domain

RT fingers/palm

Thumb/X

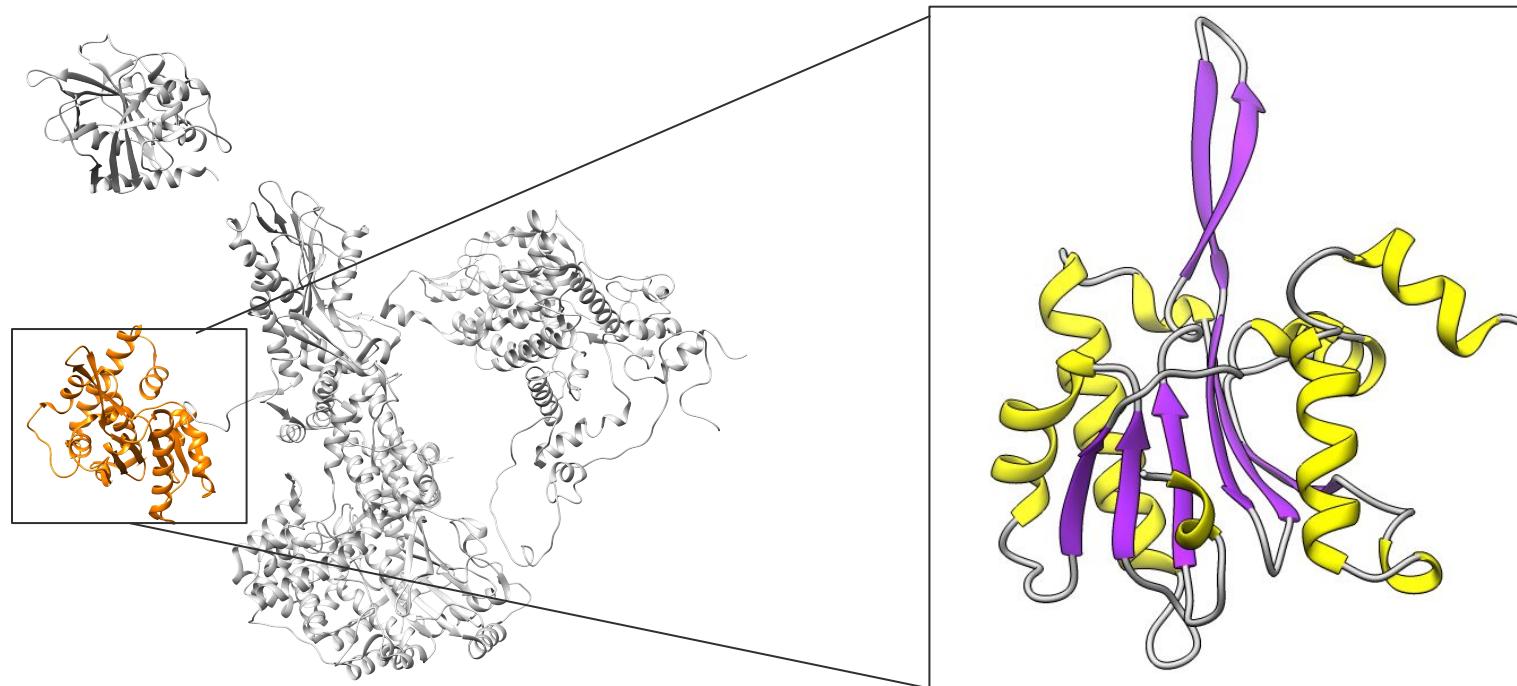
Linker

Endonuclease

RNase H-like

Jab1/MPN

RNase-H



PDB code: 4I43
Resolution: 2A

N-domain

RT fingers/palm

Thumb/X

Linker

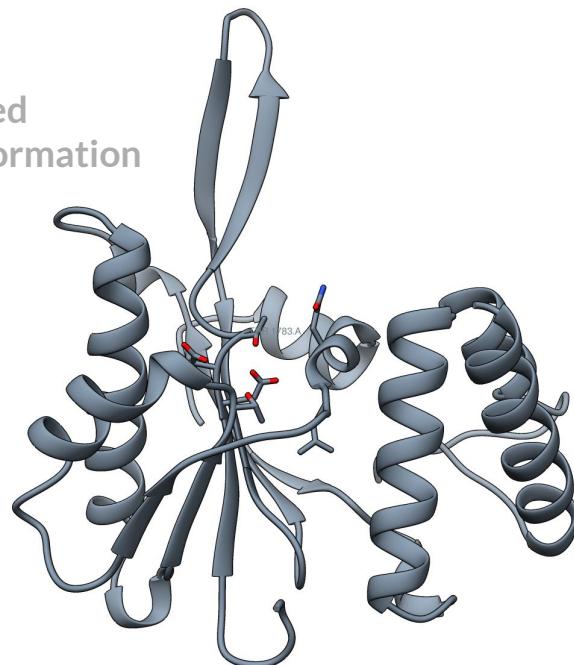
Endonuclease

RNase H-like

Jab1/MPN

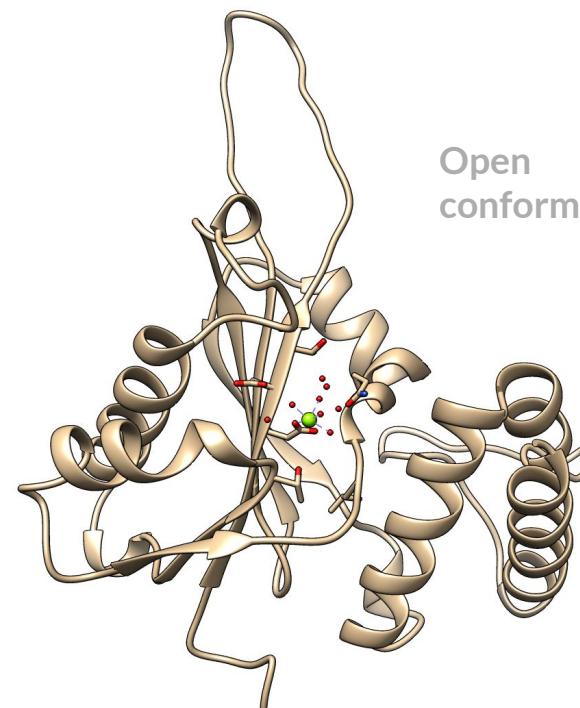
RNase-H

Closed
conformation



PDB code: 4JK7
Resolution: 1.4A

Open
conformation



N-domain

RT fingers/palm

Thumb/X

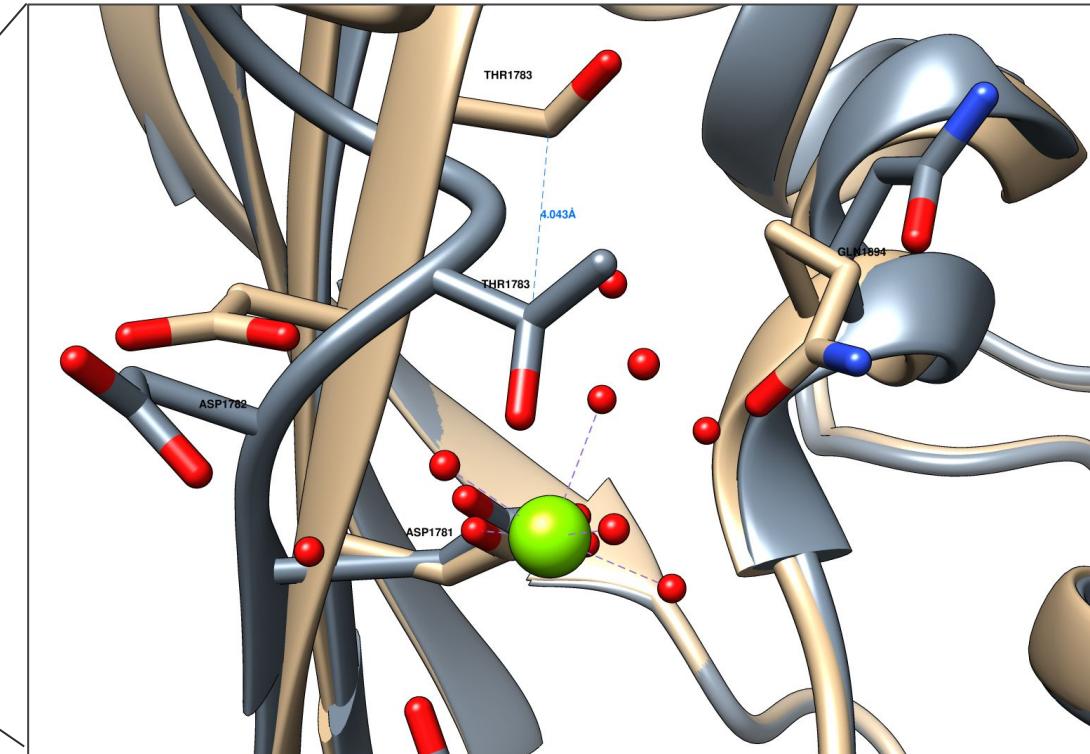
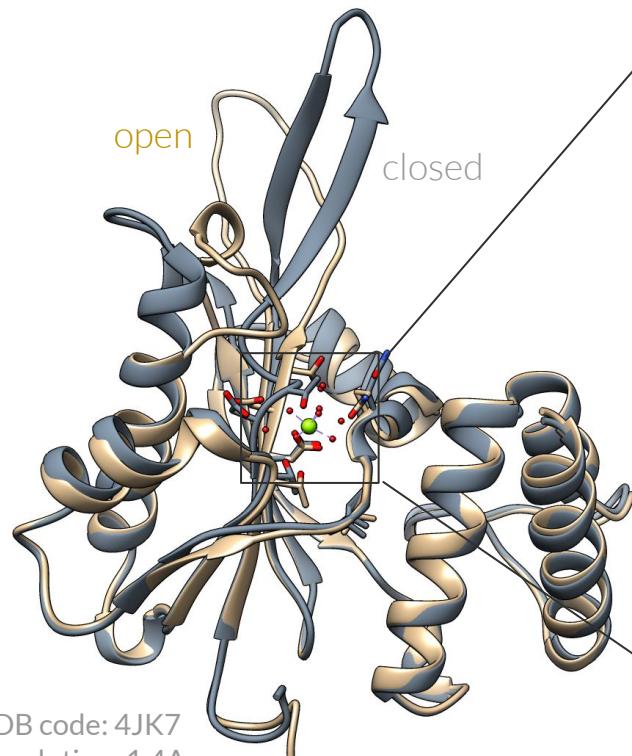
Linker

Endonuclease

RNase H-like

Jab1/MPN

RNase-H



RNase-H

		Asp 1781 Asp 1782	Tyr 1783	
	1861			1920
H. sapiens	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
C. lupus	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
M. musculus	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
M. mulatta	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
G. gallus	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
X. tropicalis	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
D. rerio	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
D. melanogaster	YLSSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
C. elegans	YLTSQNYGEL	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
A. thaliana	YLSSQNYGEI	FSNQIIWFVD	DT	NVYRVTIH KTFEGNLTTK PINGVIFIFN PRTGQLFLKI
S. pombe	YLSSSNYael	FSNQIQLFVD	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKV
N. crassa	FLNSQNYSEL	FSNQTQLFID	DT	NVYRVTIH KTFEGNLTTK PINGAIFIFN PRTGQLFLKI
S. cerevisiae	FLNSSNYael	FNNDIKLFVD	DT	NVYRVTIH KTFEGNVATK AINGCIFTLN PKTGHLFLKI

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

RNase H-like

Jab1/MPN

RNase-H

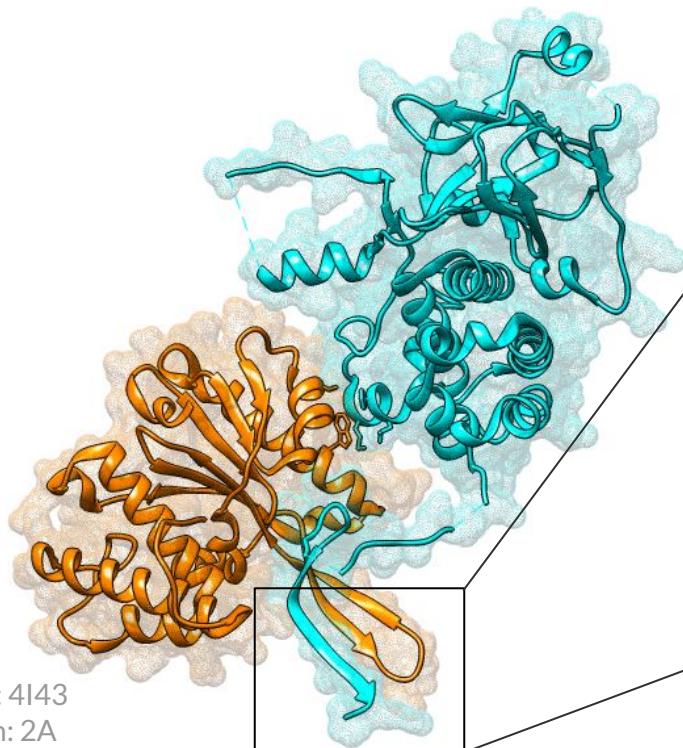


STAMP VALUES
Sc 7.11
RMS 1.24

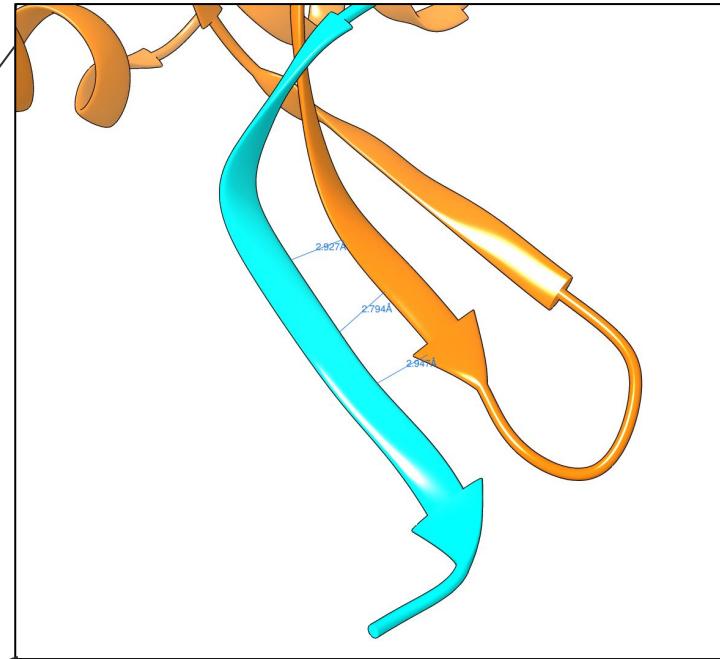
PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A

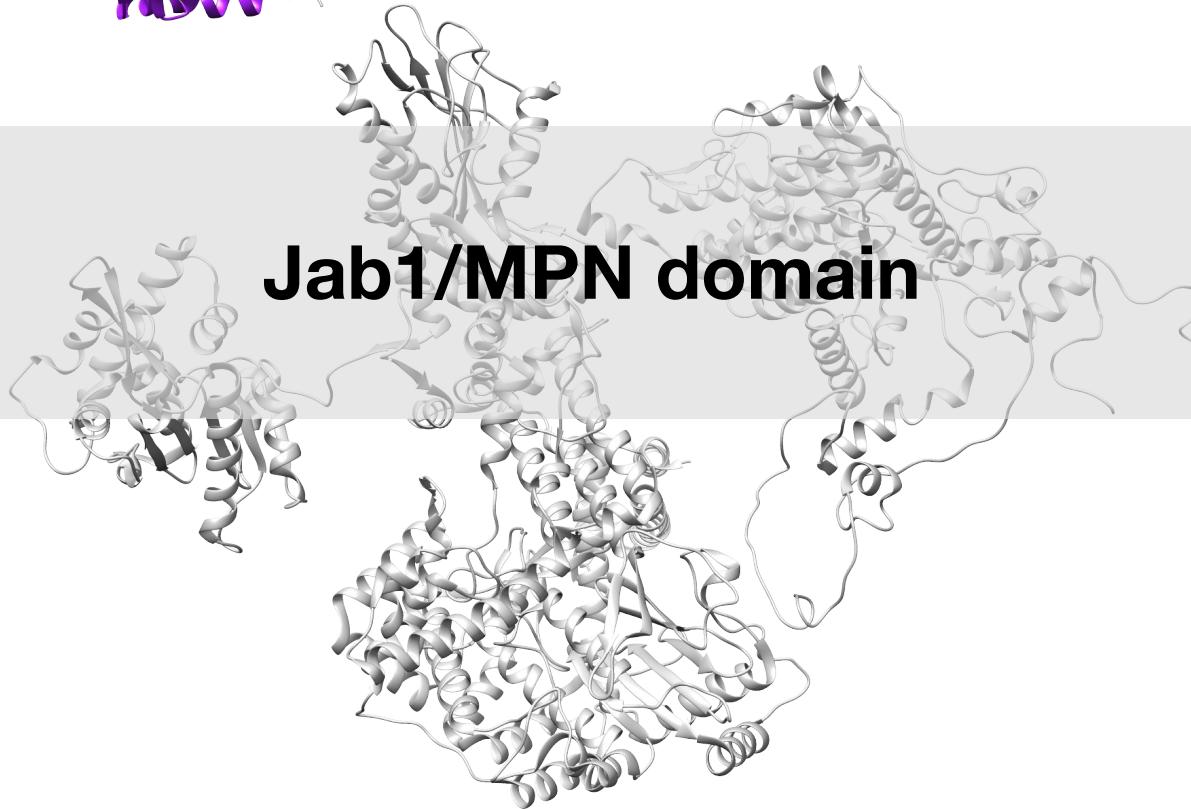
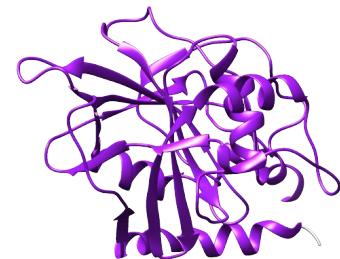
RNase-H



Aar2 stabilizes the β -hairpin...



...in order to make it stay in the pre-catalytic conformation (closed)



Jab1/MPN domain

N-domain

RT fingers/palm

Thumb/X

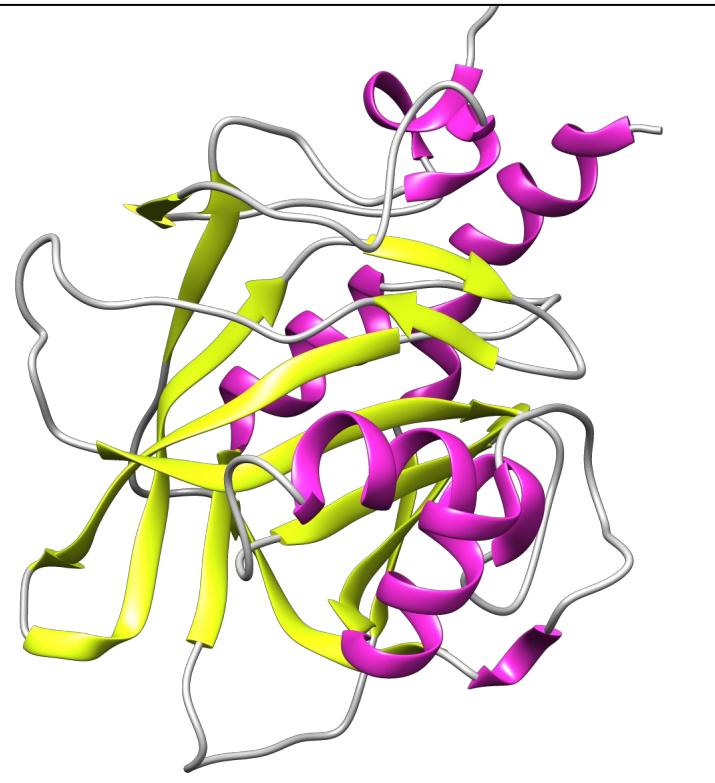
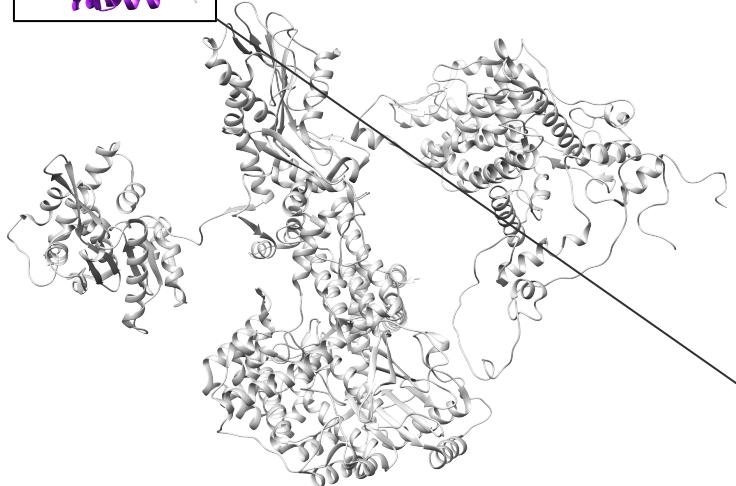
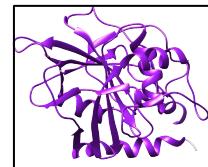
Linker

Endonuclease

RNase H-like

Jab1/MPN

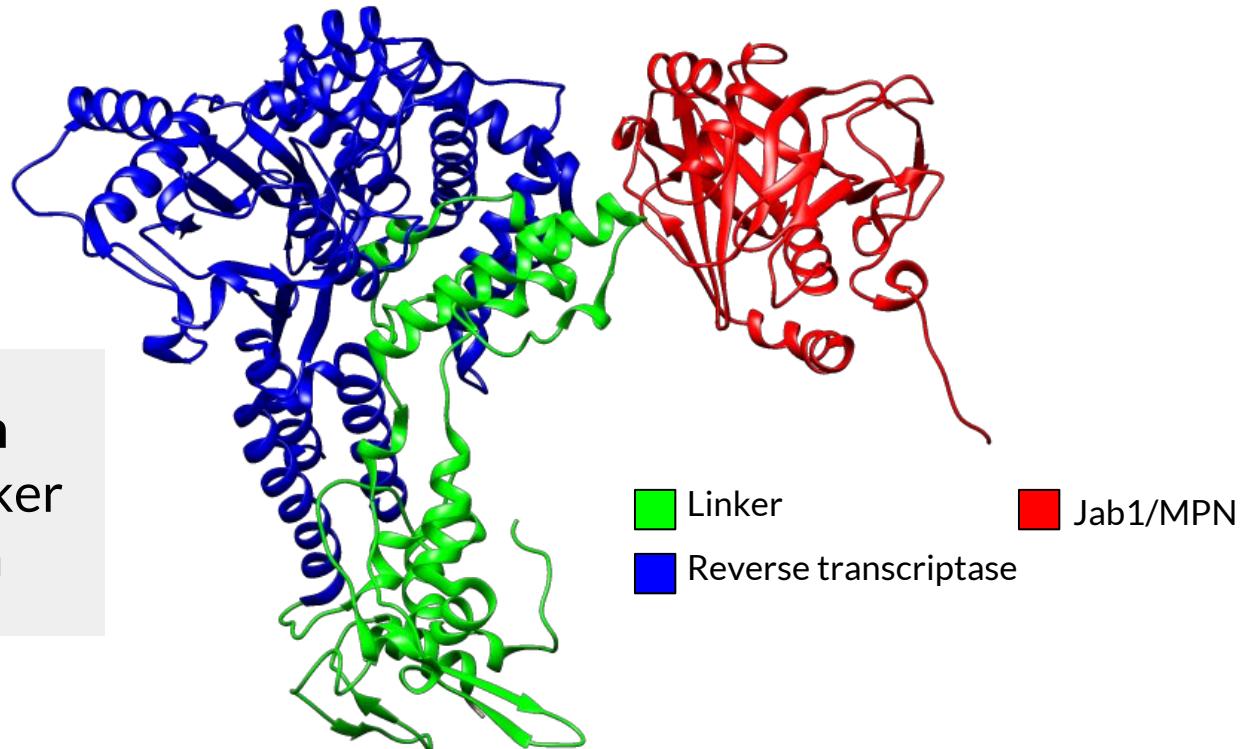
Jab1/MPN



PDB code: 4I43
Resolution: 2A

Jab1/MPN

Jab1/MPN domain
interacts with the linker
and the RT domain



N-domain

RT fingers/palm

Thumb/X

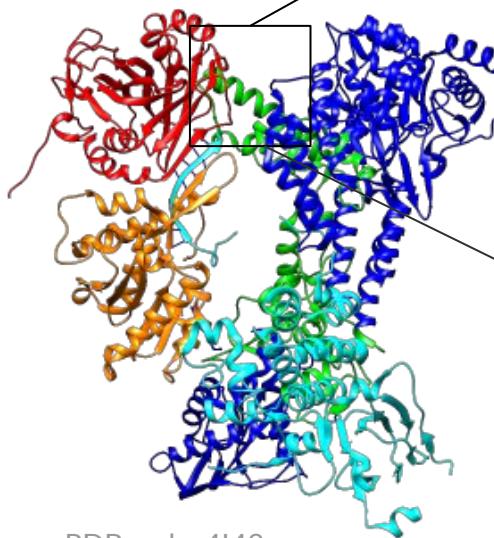
Linker

Endonuclease

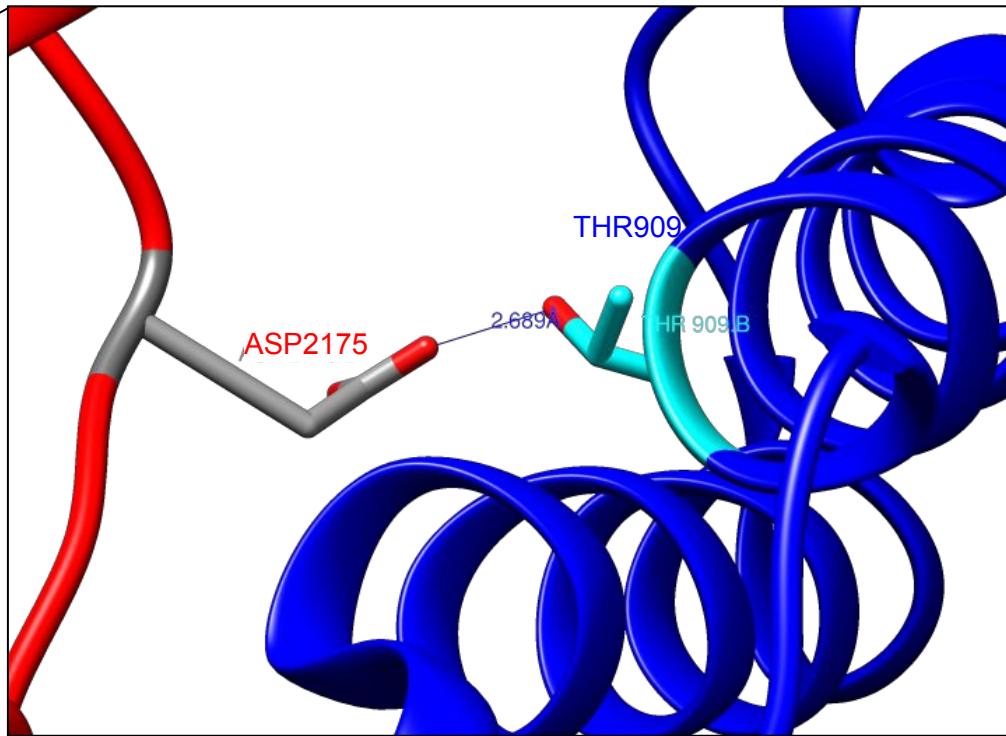
RNase H-like

Jab1/MPN

Jab1/MPN



PDB code: 4I43
Resolution: 2A



■ **Jab1/MPN**
■ **RT**

N-domain

RT fingers/palm

Thumb/X

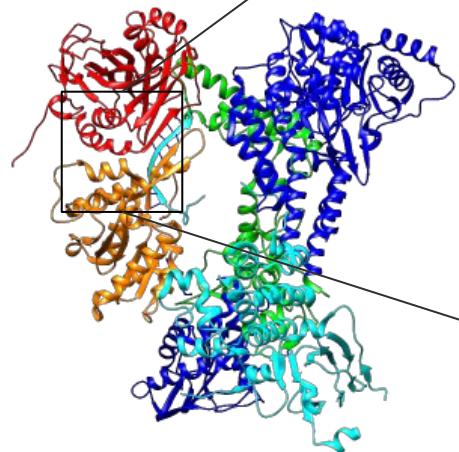
Linker

Endonuclease

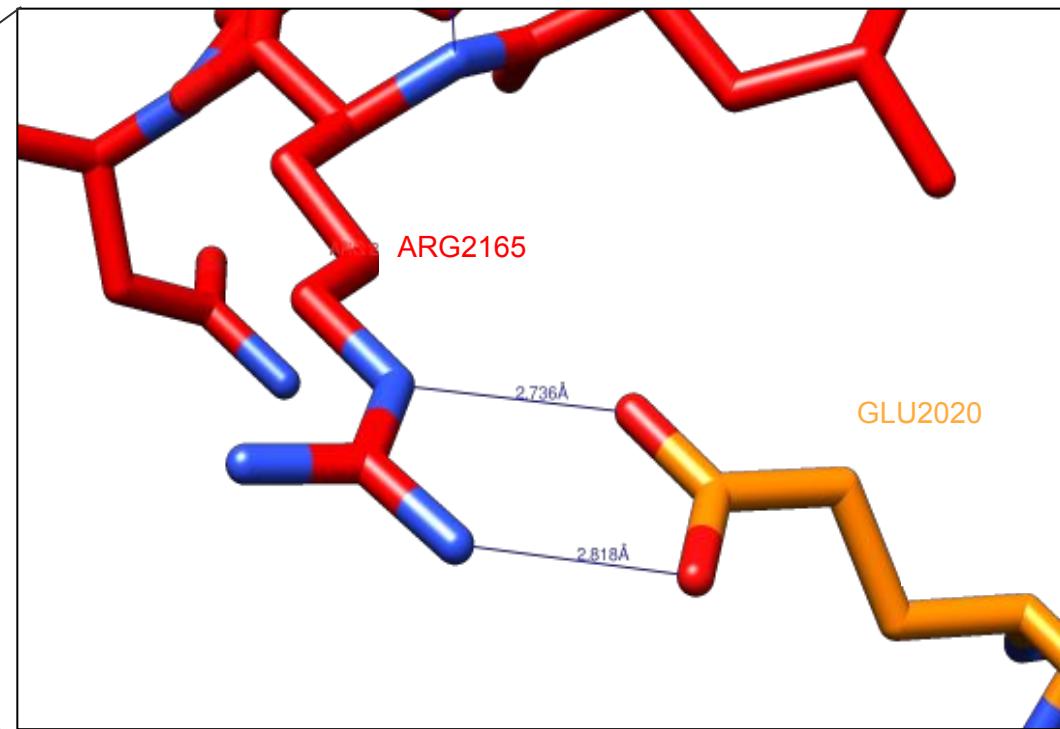
RNase H-like

Jab1/MPN

Jab1/MPN



PDB code: 4I43
Resolution: 2A



■ Jab1/MPN
■ RNaseH-like

Jab1/MPN

RNAseH-like region

Glu 2020

1981

2040

H.sapiens	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTITEPHHIW
M.mulatta	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTITEPHHIW
C.lupus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTITEPHHIW
M.musculus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTVTEPHHIW
G.gallus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTITEPHHIW
D.rerio	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTITEPHHIW
D.melanogaster	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNTE	RTKIIILKPDK	TTITEAHHIW
C.elegans	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRVVL	IMRGMHINPD	KTKVILKPDK	TTITEPHHIW
S.cerevisiae	ATEPQMVLFN	IYDDWLDRIS	SYTAFSRLTL	LLRALKTNNE	SAKMILLSDP	TITIKSYHLW
S.pombe	ATEPQMVLFN	LYDDWLQSVS	SYTAFSRLIL	ILRALNVNTE	KTKLILRPDK	SIITKENHVV
N.crassa	ATEPQMVLFN	LYDEWLKSIS	SYTAFSRLIL	ILRALHVNQD	KTKLILRPDK	TVITQDHIIW
A.thaliana	ATEPQMALFN	IYDDWLMTVS	SYTAFQRLIL	ILRALHVNNE	KAKMLLKPDPM	SVVTEPNHIW
X.tropicalis	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNNND	RAKVILKPDK	TTVTEPHHIW

Negatively charged

Non-polar

Positively charged

Polar

Sequence alignment based on PFAM domains

Jab1/MPN

RT region

Thr 909

	901	960									
H.sapiens	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
C.lupus	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
M.musculus	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
M.mulatta	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
G.gallus	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
X.tropicalis	AVAVYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
D.rerio	AVAIYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
D.melanogaster	AVAIYTTTVH	WLESRRFAPI	PFPPPLSYKHD	TKLLLILALER	LKEAYSVKSR	LNQSQREELG					
C.elegans	AVAIYTTTVH	WLESRRFSPI	PFPPPLSYKHD	TKLLLILALER	LKESYSVKNR	LNQSQREELA					
A.thaliana	GIAIYTTTVN	WLESRKFSAI	PFPPPLSYKHD	TKLLLILALER	LKESYSAAVK	LNQQQREELG					
S.pombe	AVAIYTTFVH	WLESRRFQPI	PFPPPLSYKHD	TKLLVLALER	LKEAYSVKGR	LNQSQREELA					
N.crassa	AVAIYTTTVH	WLESRKFSPI	PFPSVSYKHD	TKILILALER	LREAYSTKGR	LNQSQREELA					
S.cerevisiae	ATTIFSVVMVE	WLESRSFSPI	PFPPPLTYKND	TKILVLALED	LKDVYASKVR	LNASEREELA					

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

Jab1/MPN

Jab1/MPN region

Arg 2165

Asp 2175

	2161											2220
<i>H. sapiens</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>C. lupus</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>M. musculus</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>M. mulatta</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>G. gallus</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>X. tropicalis</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>D. rerio</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISAANLH
<i>D. melanogaster</i>	ITSTTSNYET	Q	T	F	S	S	K	T	E	W	R	VRAISATNLH
<i>C. elegans</i>	ITATTTSNYET	A	S	F	A	S	R	T	E	W	R	VRAISSTNLH
<i>A. thaliana</i>	ISTTISPYEQ	S	A	F	G	S	K	T	D	W	R	VRAISATNLY
<i>S. pombe</i>	VVTTTSAYEN	E	K	F	S	S	K	T	E	W	R	NRAISSISLP
<i>N. crassa</i>	IVTTTSQFEQ	Q	T	F	A	S	K	T	E	W	R	TRAIATSNLR
<i>S. cerevisiae</i>	VVVVASADYES	Q	T	F	S	S	K	N	E	W	R	KAIAINTLLY

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

N-domain

RT fingers/palm

Thumb/X

Linker

Endonuclease

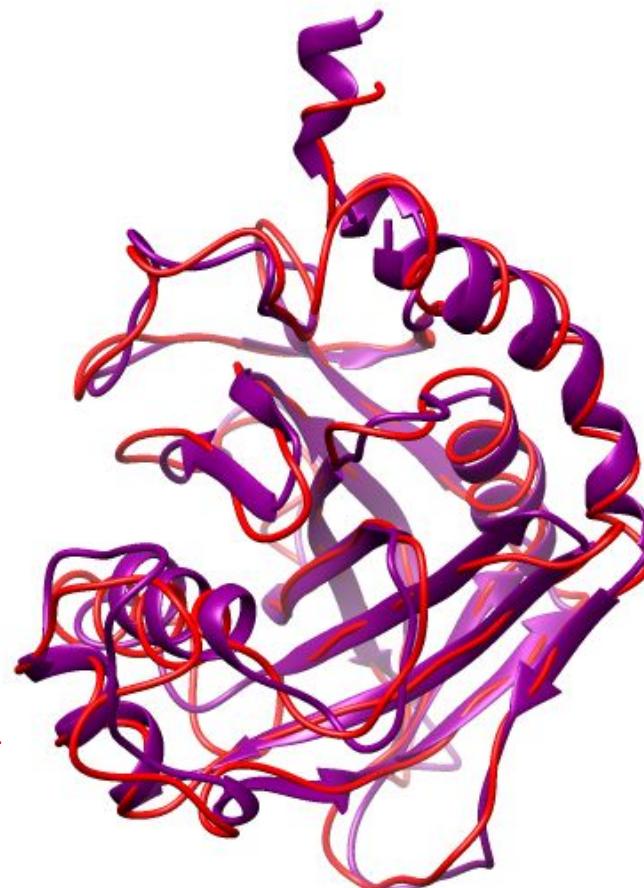
RNase H-like

Jab1/MPN

Jab1/MPN

yPrp8 Jab1

hPrp8 Jab1



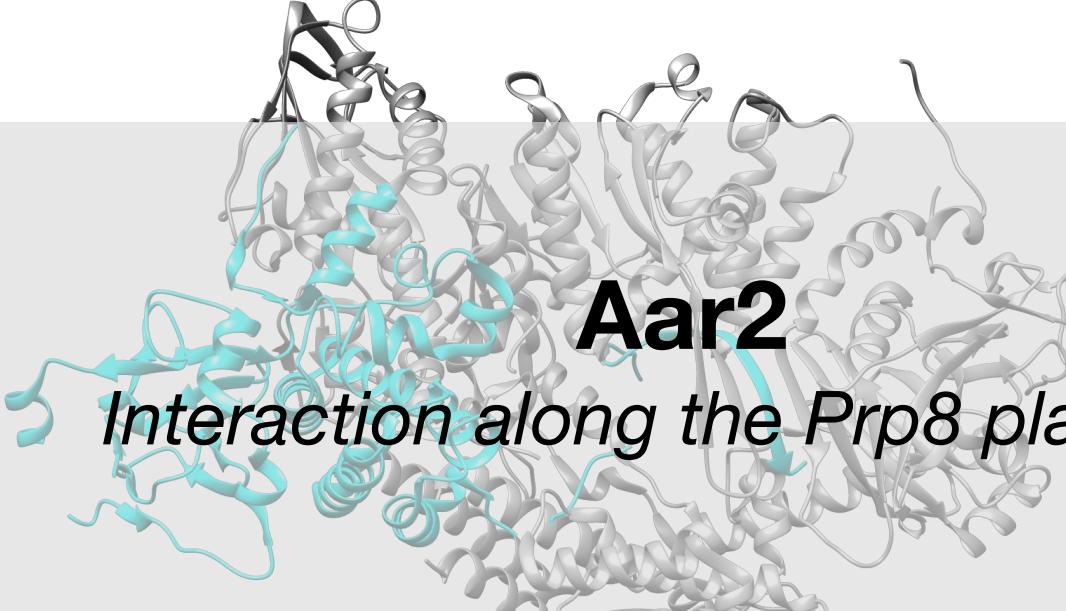
STAMP VALUES

Sc 8,02

RMS 1.29

PDB code: 3JCM
Resolution: 3.8A

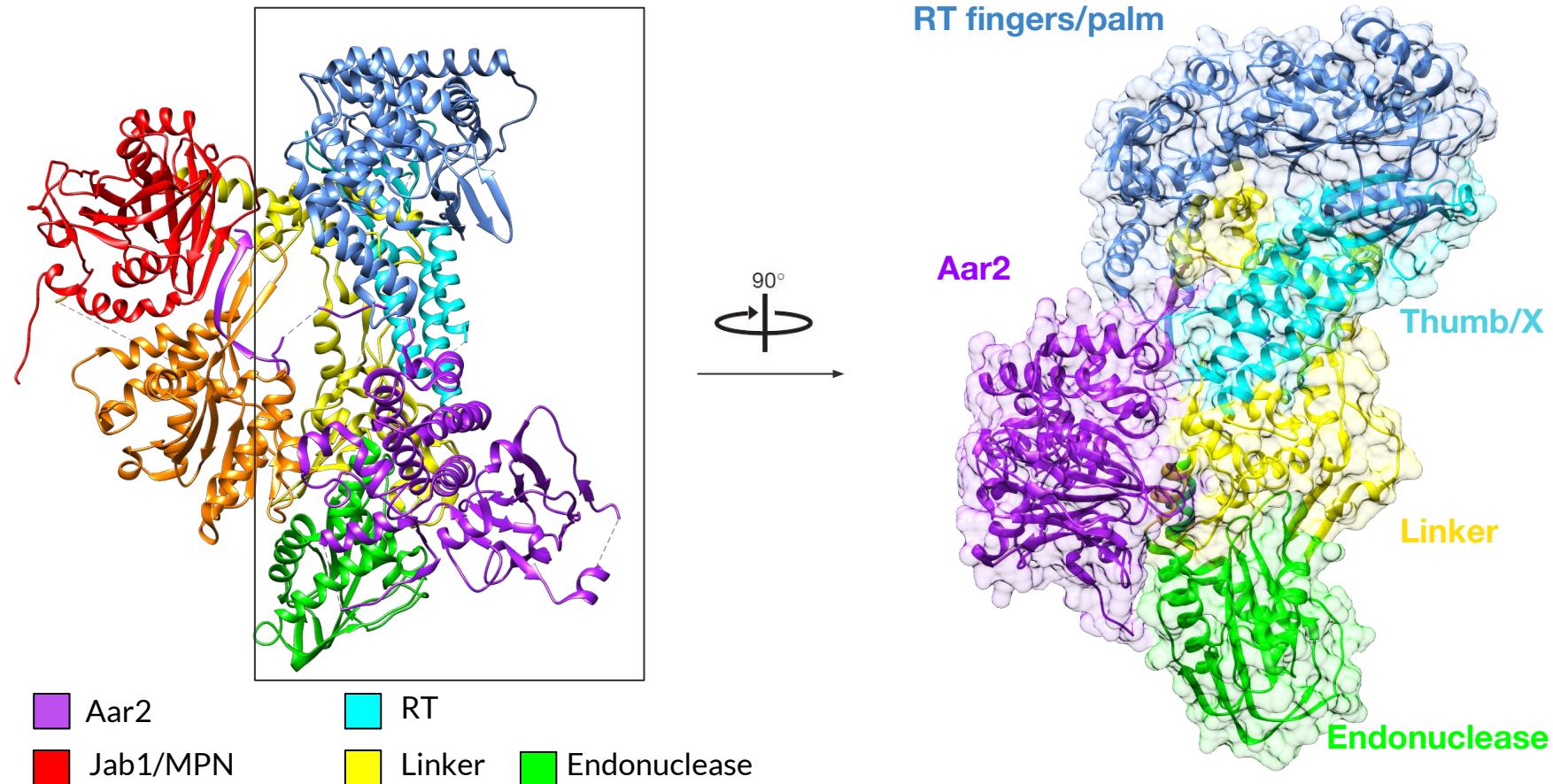
PDB code: 3JCR
Resolution: 7A



Aar2

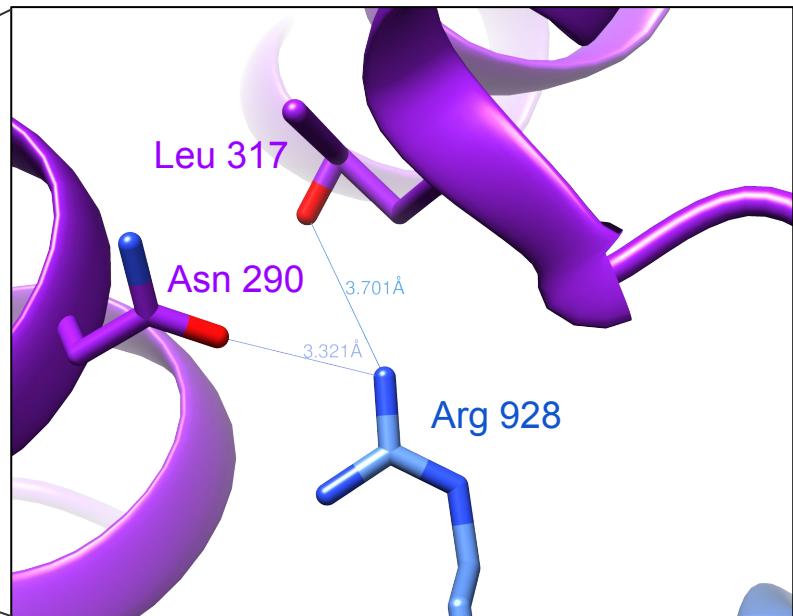
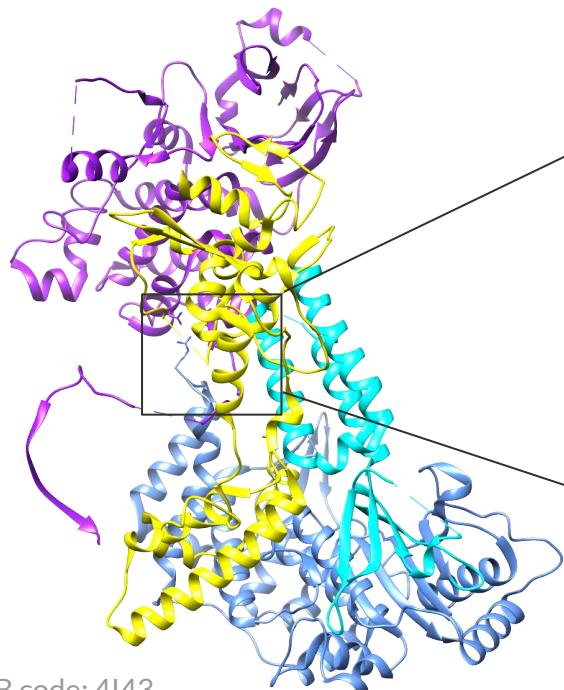
Interaction along the Prp8 platform

Aar2



Aar2

Aar2 - RT



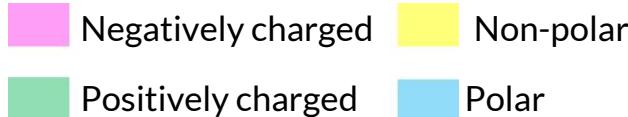
■ RT fingers/palm

■ Aar2

Aar2

Aar2 - RT

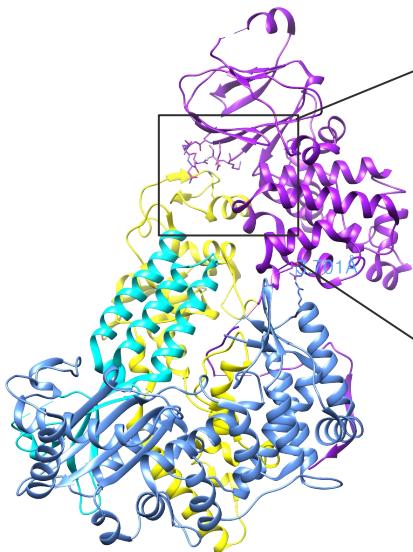
	<u>RT region</u>						Arg 928
	901						960
H. sapiens	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
C. lupus	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
M. musculus	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
M. mulatta	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
G. gallus	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
X. tropicalis	AVAVYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
D. rerio	AVAIYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
D. melanogaster	AVAIYTTTVH	WLESRRFAPI	PFPPLSYKHD	TKLLILALER	LKEAYSV K SR	LNQSQREELG	
C. elegans	AVAIYTTTVH	WLESRRFSPI	PFPPLSYKHD	TKLLILALER	LKESYSV K NR	LNQSQREELA	
A. thaliana	GIAIYSTTVN	WLESRKFSAI	PFPPLSYKHD	TKLLILALER	LKESYSA A VK	LNQQQREELG	
S. pombe	AVAIYTTFVH	WLESRRFQPI	PFPPLSYKHD	TKLLVLAER	LKEAYSV K GR	LNQSQREELA	
N. crassa	AVAIYTTTVH	WLESRKFSPI	PFPVSYSYKHD	TKILILALER	LREAYST K GR	LNQSQREELA	
S. cerevisiae	ATTIFSVVMVE	WLESRSFSPI	PFPPLTYKND	TKILVLAED	LKDVYASK V R	LNASEREELA	



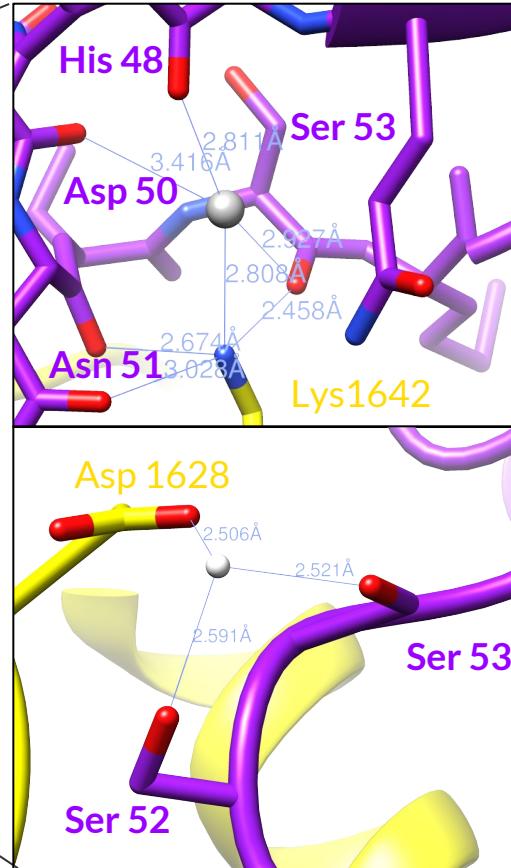
Clustal Sequence alignment

Aar2

Aar2 - Linker



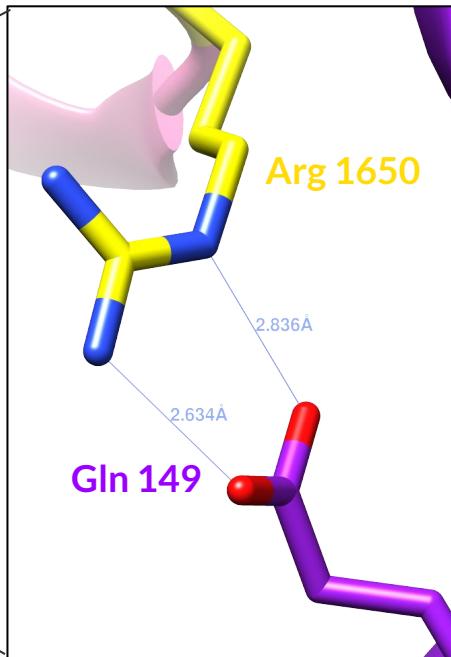
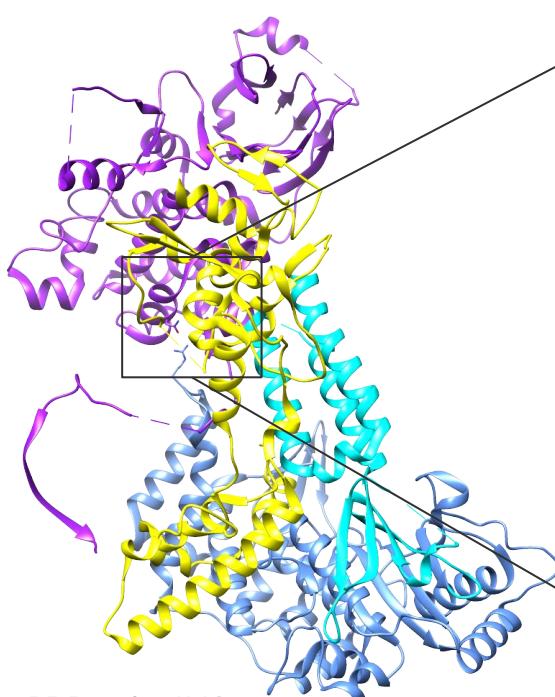
PDB code: 4I43
Resolution: 2A



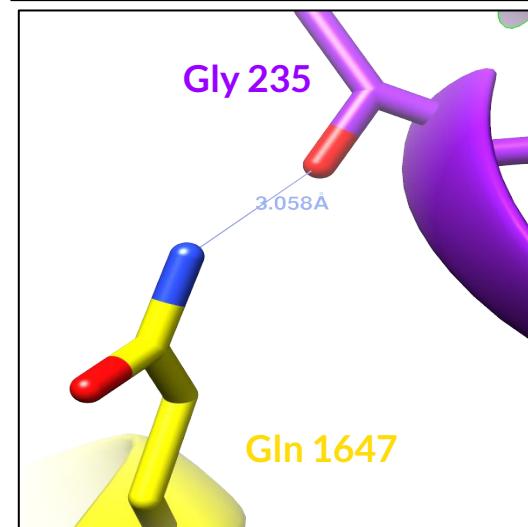
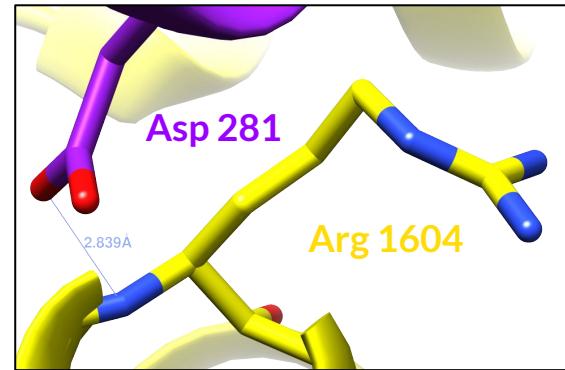
Aar2
Linker

Aar2

Aar2 - Linker



■ Aar2 ■ Linker



Aar2

Aar2 - Linker

Linker region

Arg 1604

1620

	1561	<u>Linker region</u>					Arg 1604	1620
H. sapiens	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
M. mulatta	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
C. lupus	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
M. musculus	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
G. gallus	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
D. rerio	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
D. melanogaster	LFKGTYFPTW	EGLFWEKASG	FEESMKYKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
C. elegans	LFRGTYFPTW	EGLFWERASG	FEESMKFKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
S. cerevisiae	LFKGTFNSW	EGLFWEKASG	FEDSMQFKKL	THAQRTGLSQ	IPNRRFTLWW	SPTINRANVY		
S. pombe	MFKATGFPFW	EGLFWEKASG	FEESMKFKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
N. crassa	LFKATGFPFW	EGLFWEKASG	FEESMKFKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
A. thaliana	LFKGTYFPTW	EGLFWEKASG	FEESMKYKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		
X. tropicalis	LFKGTYFPTW	EGLFWEKASG	FEESMKWKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY		

Negatively charged

Non-polar

Positively charged

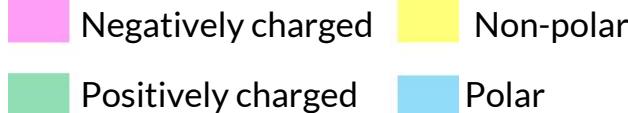
Polar

Sequence alignment based on PFAM domains

Aar2

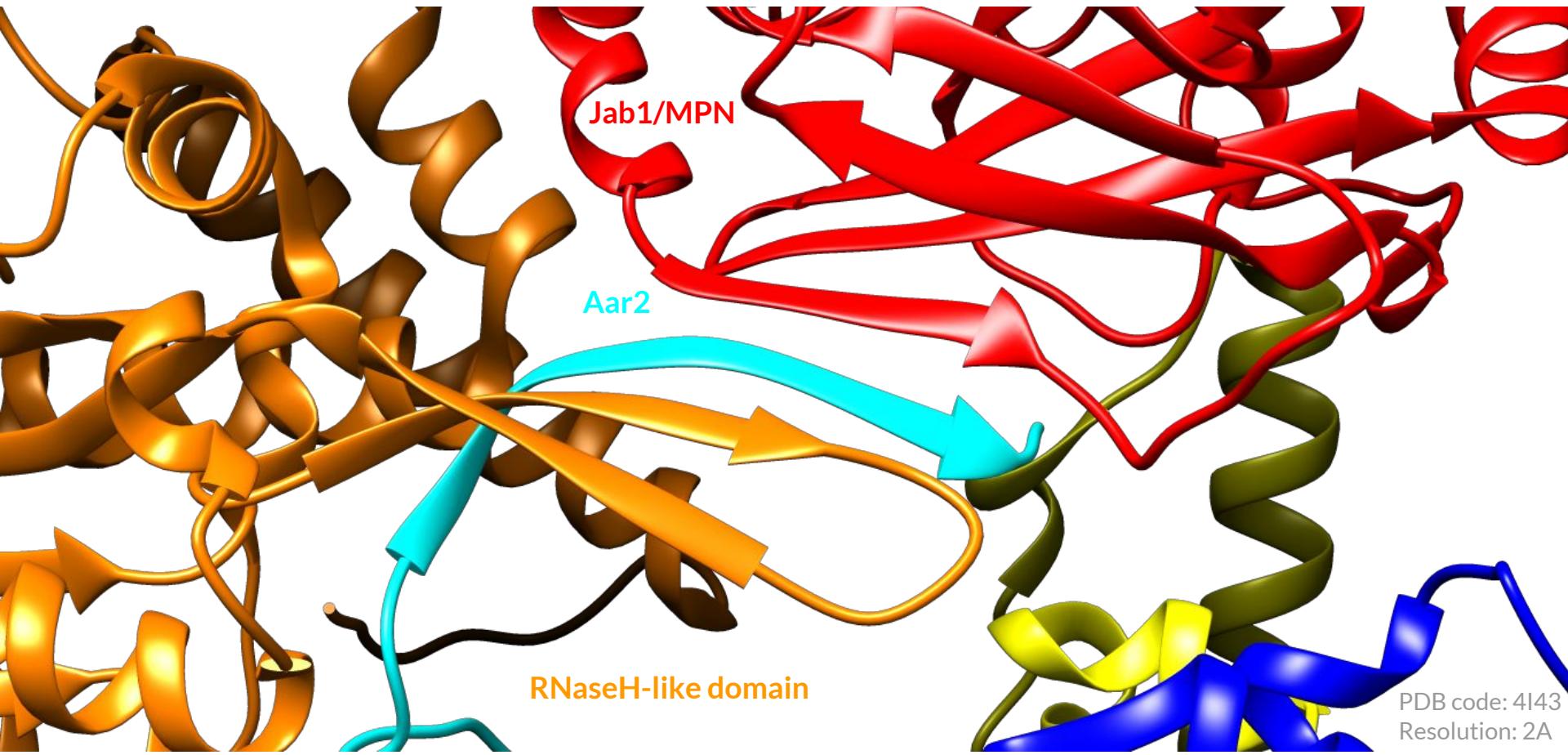
Aar2 - Linker

		Asp 1628	Lys 1690	Gln 1647	Arg 1650	Linker region	
	1621					1680	
H. sapiens		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
M. mulatta		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
C. lupus		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
M. musculus		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
G. gallus		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
D. rerio		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
D. melanogaster		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet
C. elegans		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	VVMDLCQVFD	QELdaleiqt
S. cerevisiae		VGFLVQLDLT	GIFLGKIP	L KISLIQIFR	AHLWQKIHES	IVFDICQILD	GELdvlqies
S. pombe		VGFQVQLDLT	GIMMHGKIPT	L KISLIQIFR	SHLWQKIHES	VVWDLCQVLD	QELeislqiet
N. crassa		VGFQVQLDLT	GIFLGKIP	L KISLIQIFR	AHLWQKIHES	VVMDLCQVFD	QELealsies
A. thaliana		VGFQVQLDLT	GIYMHGKIPT	L KISLIQIFR	AHLWQKIHES	VVMDLCQVLD	QELepleiet
X. tropicalis		VGFQVQLDLT	GIFMHGKIPT	L KISLIQIFR	AHLWQKIHES	IVMDLCQVFD	QELdaleiet



Sequence alignment based on PFAM domains

Aar2



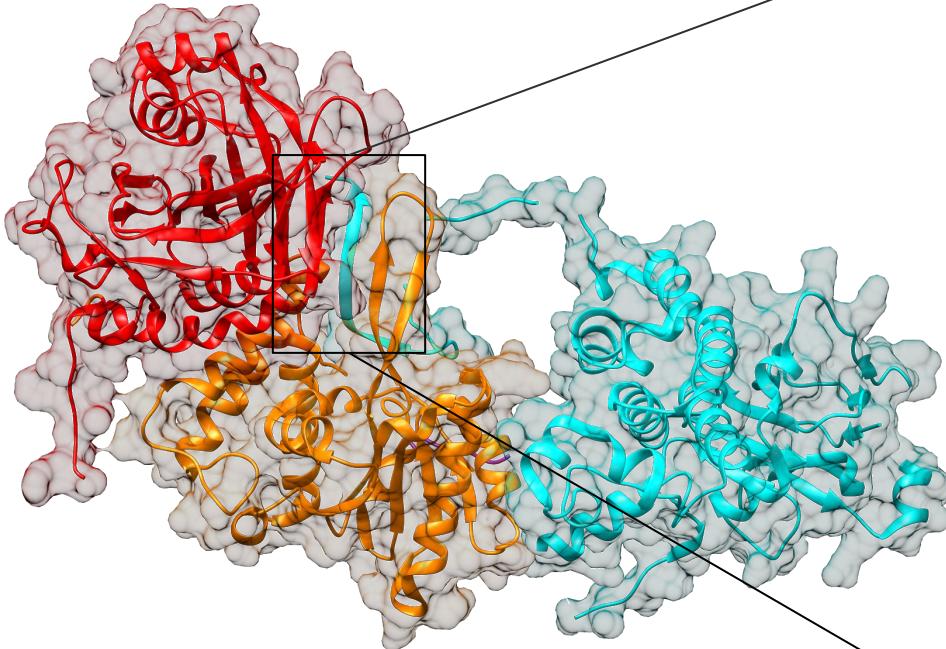
RNaseH-like domain

PDB code: 4I43

Resolution: 2A

Aar2

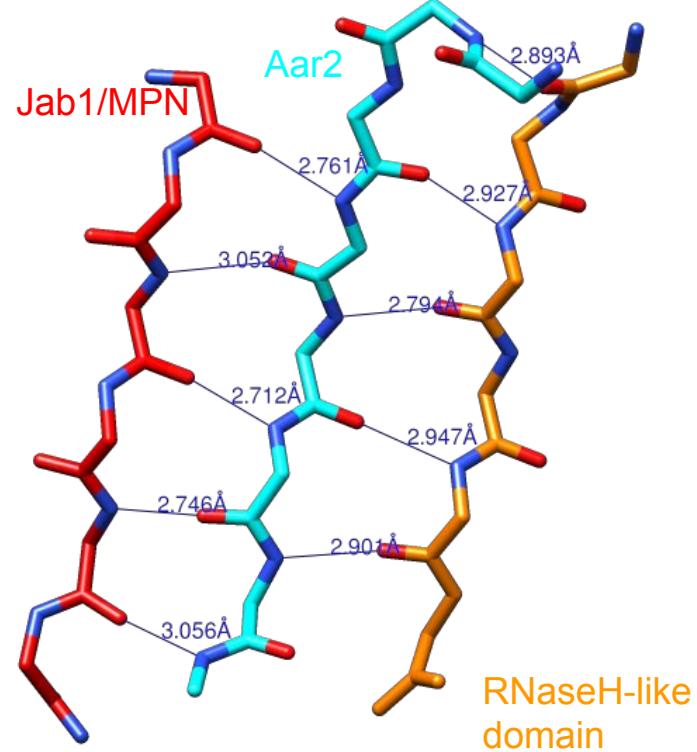
Aar2 - Jab1/RNase



PDB code: 4I43
Resolution: 2A

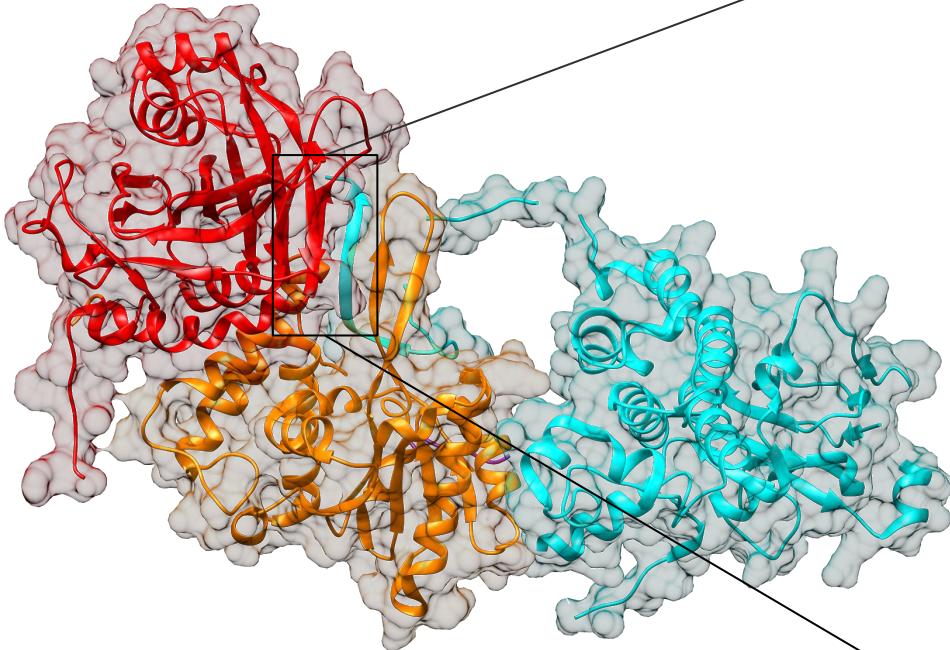
■ RNaseH-like
■ Jab1/MPN

■ Aar2



Aar2

Aar2 - Jab1/RNase



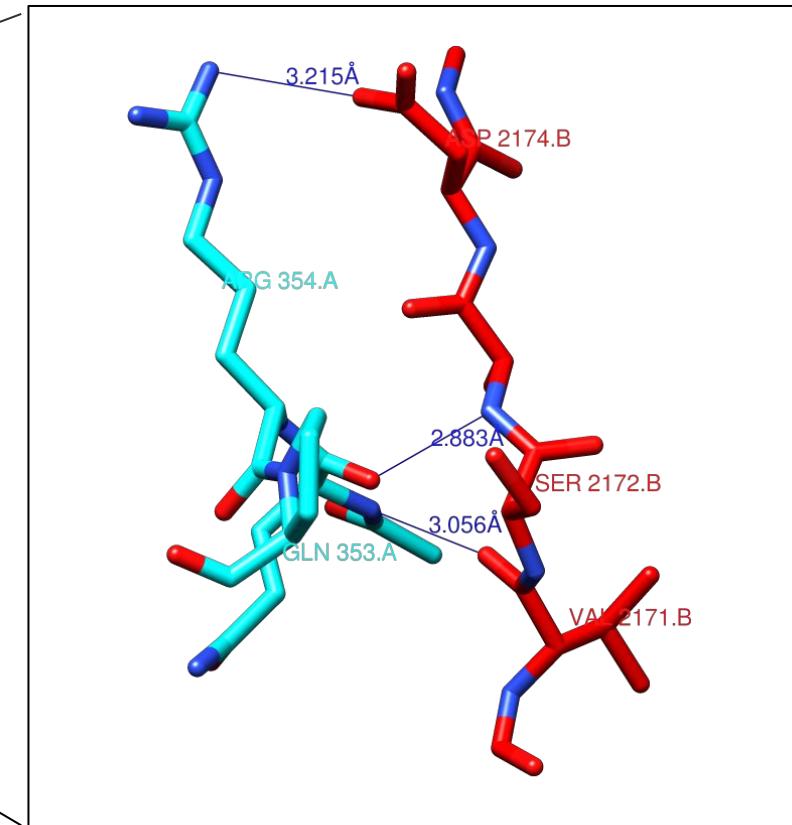
RNaseH-like



Jab1/MPN

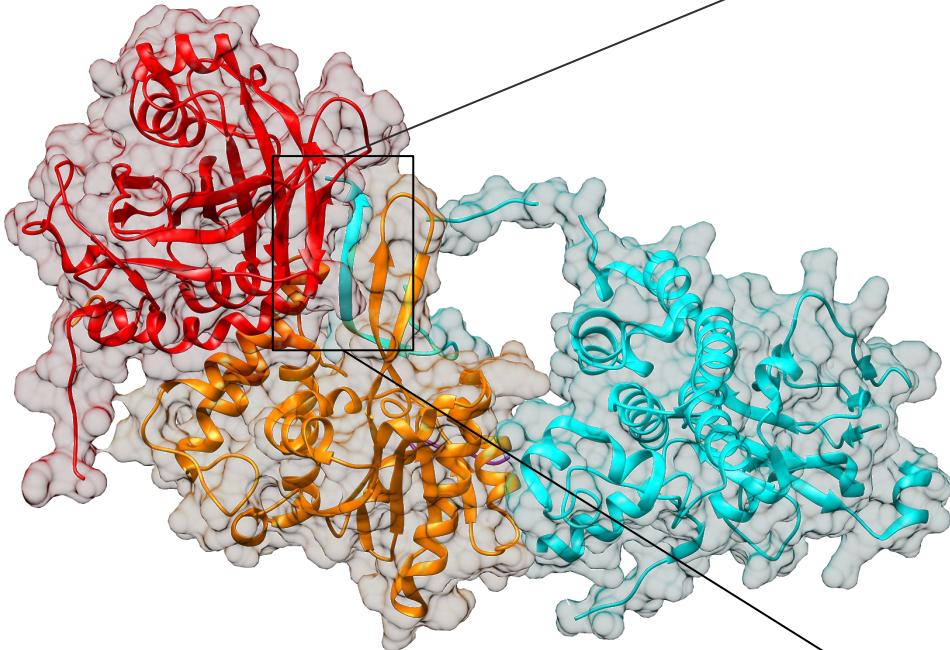


Aar2



Aar2

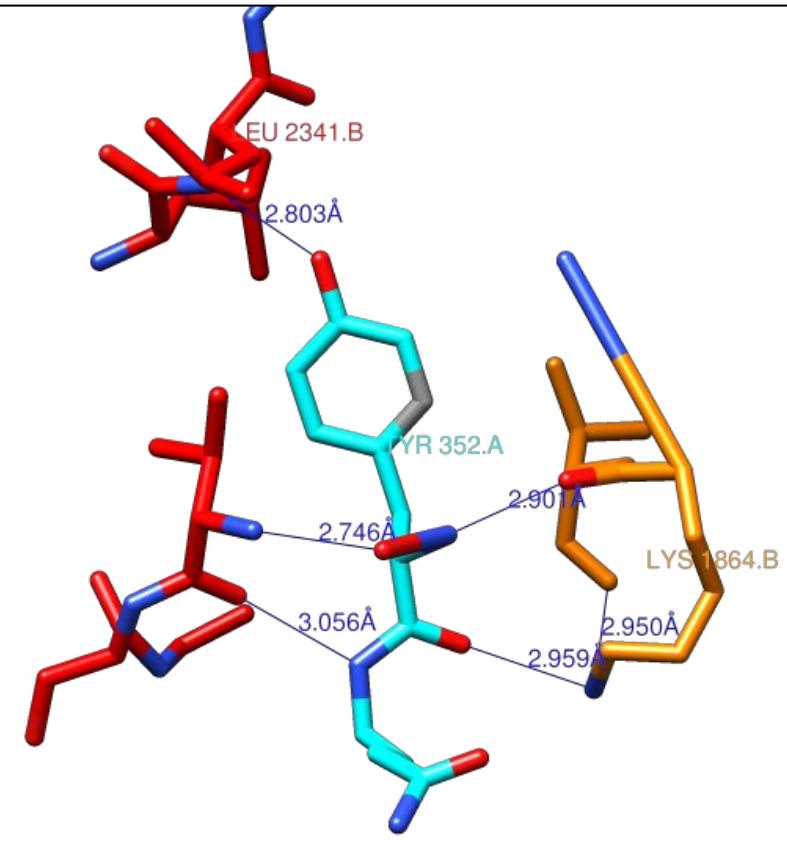
Aar2 - Jab1/RNase



PDB code: 4I43
Resolution: 2A

■ RNaseH-like
■ Jab1/MPN

■ Aar2



Aar2

	Leu 2341	<u>Jab1/MPN region</u>				
	2341					2400
H. sapiens	LSDRFLGFFM	VPAQSSWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFA
M. mulatta	LSDRFLGFFM	VPAQSSWYN	FMGK.....
C. lupus	LSDRFLGFFM	VPAQSSWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFA
M. musculus	LSDRFLGFFM	VPAQSSWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFA
G. gallus	LSDRFLGFFM	VPAQGSWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFA
D. rerio	LSDRFLGFFM	VPGQVSWYN	FMGVRHD..P	NMKYDLQLAN	PKEFYHEVHR	PSHFLNFA
D. melanogaster	LSNKFLGFFM	VPAQSSWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHELHR	TSHFLLFSNL
C. elegans	LSDRFLGYFM	VPSNGVWYN	FQGQRWS..P	AMKFDVCLSN	PKEYYHEDHR	PVHFHNFKAF
S. cerevisiae	LSDRITGNFI	IPSGNVWNY	FMGTAFN..Q	EGDYNFKYGI	PLEFYNEMHR	PVHFLQFSEL
S. pombe	LSDRIQGFFL	VPEEGVWYN	FNGASFS..P	KMTYSLKLDV	PLPFFALEHR	PTHVISYTEL
N. crassa	LSEKFRGFFL	VPDGGKWNS	FMGSAFGglE	KKPVHVVKLDT	PLPFYSDQHR	PIHFSSFNEL
A. thaliana	LSDRFFGFFM	VPENGPNYN	FMGANHT..V	SINYSLTLGT	PKEYYHQVHR	PTHFLQFSKM
X. tropicalis	LSDRFLGFFM	VPAQASWYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFA

Negatively charged

Non-polar

Positively charged

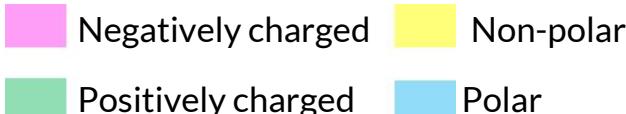
Polar

Sequence alignment based on PFAM domains

Aar2

Jab1/MPN region

	2161	Val 2171	Ser 2172	2220
<i>H. sapiens</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>C. lupus</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>M. musculus</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>M. mulatta</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>G. gallus</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>X. tropicalis</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>D. rerio</i>	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK			
<i>D. melanogaster</i>	ITSTTSNYET QTFSSKTEWR VRAISATNLH LRTNHIYVSS DDIKETG.YT YILPKNILKK			
<i>C. elegans</i>	ITATTNSNYET ASFASRTEWR VRAISSTNLH LRTQHIYVNS DDVKDTG.YT YILPKNILKK			
<i>A. thaliana</i>	ISTTISPYEQ SAFGSKTDWR VRAISATNLY LRVNHIYVNS DDIKETG.YT YIMPKNILKK			
<i>S. pombe</i>	VVTTTSAYEN EKFSSKTEWR NRAISSISLP LRTKNIYVNS DNISETFPYT YILPQNLLRK			
<i>N. crassa</i>	IVTTTSQFEQ QTFAASKTEWR TRAIATSNLR TRANINMYVSP VDSLDD.VT YVMPKNILKR			
<i>S. cerevisiae</i>	VVVASADYES QTFSSKNEWR KSAIANTLLY LRLKNIYVSA DDFVEEQ.NV YVLPKNLLKK			



Clustal Sequence alignment

Aar2

		Lys 1864		<u>Jab1/MPN region</u>		
	1861					1920
H.sapiens	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
C.lupus	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
M.musculus	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
M.mulatta	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
G.gallus	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
X.tropicalis	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
D.rerio	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
D.melanogaster	YLSSQNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
C.elegans	YLTSQLNYGEL	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
A.thaliana	YLSSQNYGEI	FSNQIIWFVD	DTNVYRVTIH	KTFEGNLTTK	PINGVIFIFN	PRTGQLFLKI
S.pombe	YLSSSNYAEI	FSNQIQLFVD	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKV
N.crassa	FLNSQNYSEL	FSNQTQLFID	DTNVYRVTIH	KTFEGNLTTK	PINGAIFIFN	PRTGQLFLKI
S.cerevisiae	FLNSSNYAEI	FNNDIKLFVD	DTNVYRVTVH	KTFEGNVATK	AINGCIFTLN	PKTGHLFLKI

1921

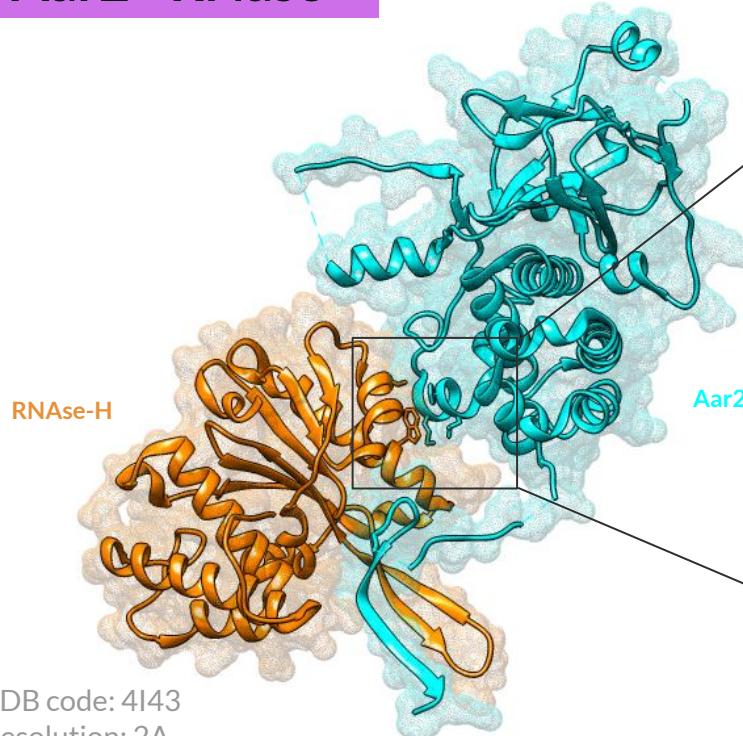
1980

Negatively charged Non-polar
Positively charged Polar

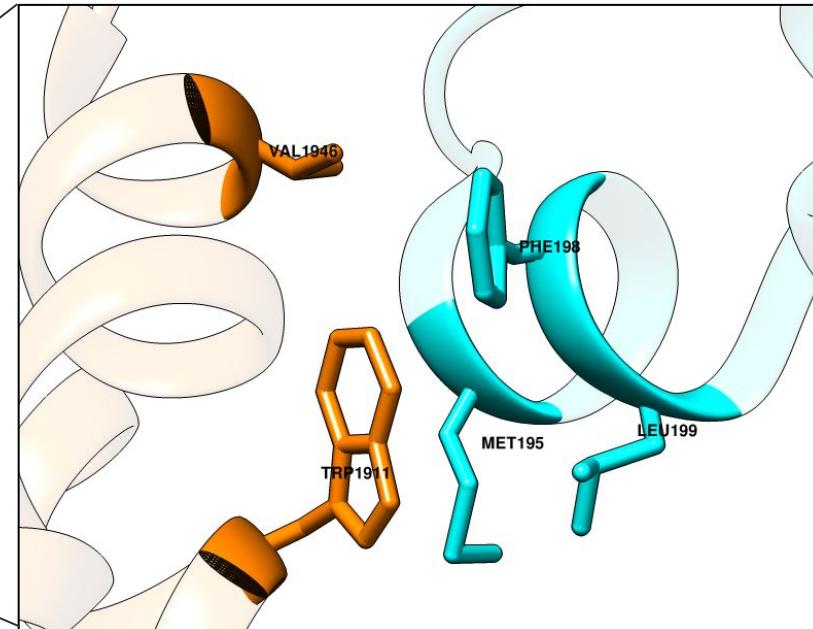
Clustal Sequence alignment

Aar2

Aar2 - RNase



Interaction of the hydrophobic residues



PDB code: 4I43
Resolution: 2A

Aar2

Trp1911

Val1946

RNaseH-like region

	1921							
H.sapiens	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
M.mulatta	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
C.lupus	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
M.musculus	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
G.gallus	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
D.rerio	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
D.melanogaster	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
C.elegans	WKTAAEVAAL	IRSLPVVEEQP	RQIIVTRKAM	LDPLEVHLLD	FPNIVIKGSE	LMLPFQAIMK		
S.cerevisiae	WKTAAEV SAL	VRSLPKEEQP	KQIIVTRKAM	LDPLEVHMLD	FPNIAIRPTE	RLLPFSAAMS		
S.pombe	WKTAAEVAAL	IRSLPVVEEQP	RQIIVTRKGM	LDPLEVHLLD	FPNITIKGSE	LQLPFQAIK		
N.crassa	WKTAAEVAAL	IRSLPVVEEQP	KQLIVTRKGL	LDPLEVNLLD	FPNISIRASE	LQLPFQAAMK		
A.thaliana	WKTAAEVAAL	VRSLPVVEEQP	KQVIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		
X.tropicalis	WKTAAEVAAL	IRSLPVVEEQP	KQIIVTRKGM	LDPLEVHLLD	FPNIVIKGSE	LQLPFQACLK		

Negatively charged

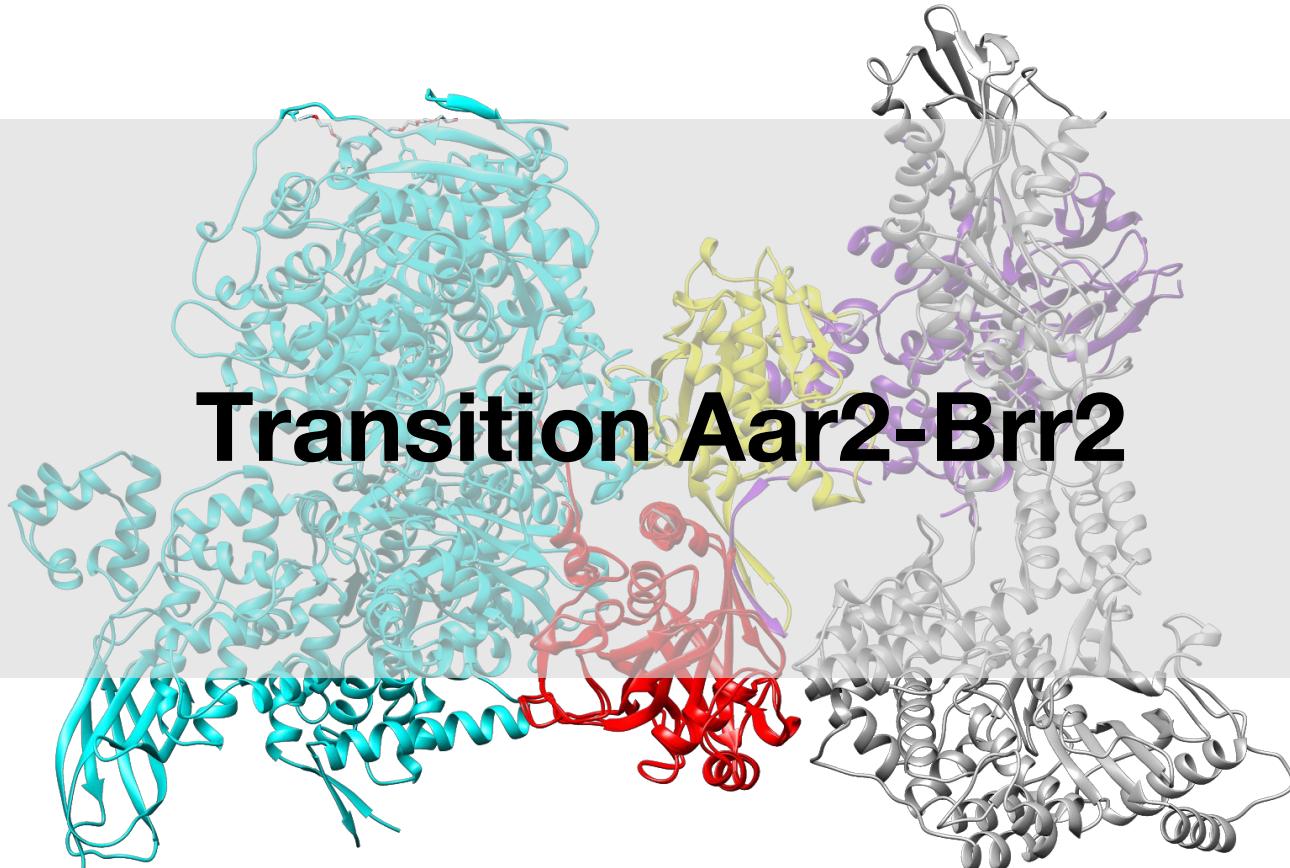
Non-polar

Positively charged

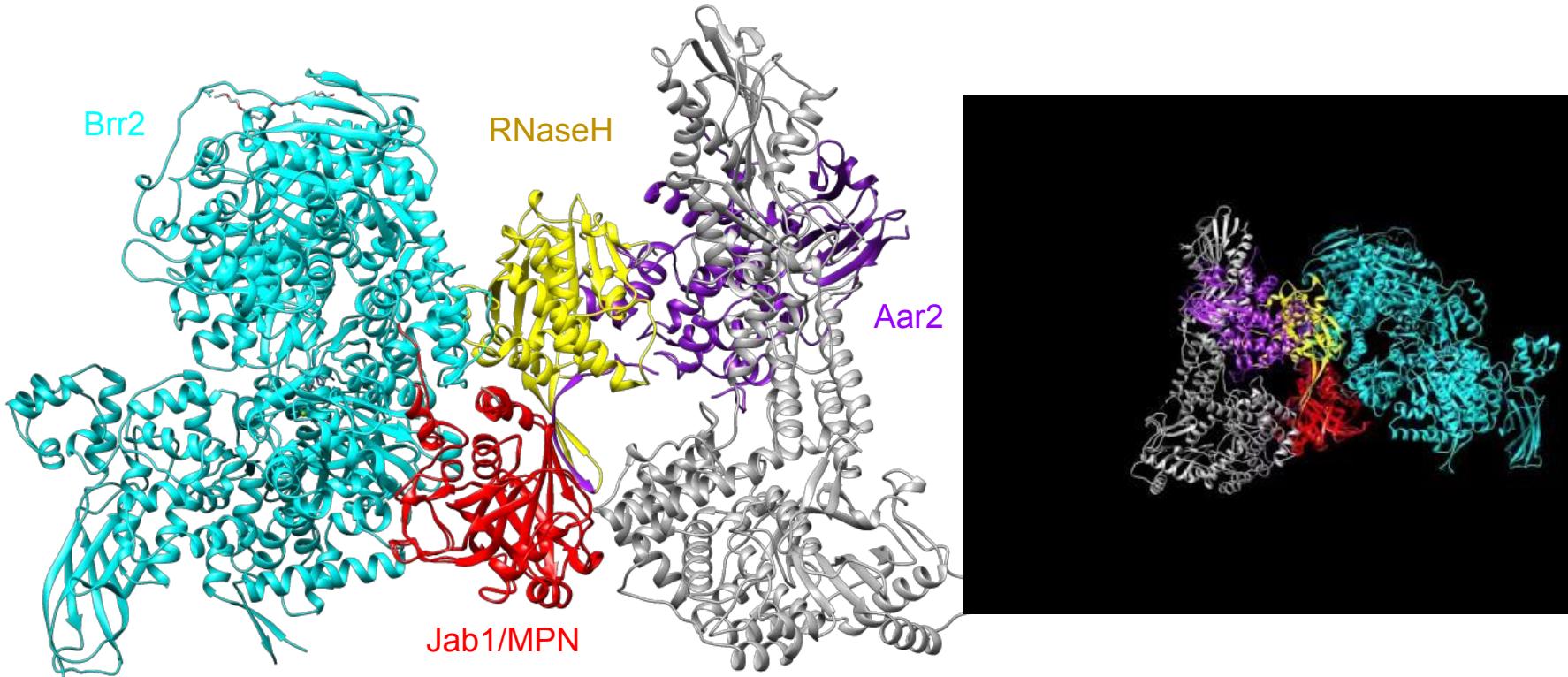
Polar

Sequence alignment based on PFAM domains

Transition Aar2-Brr2

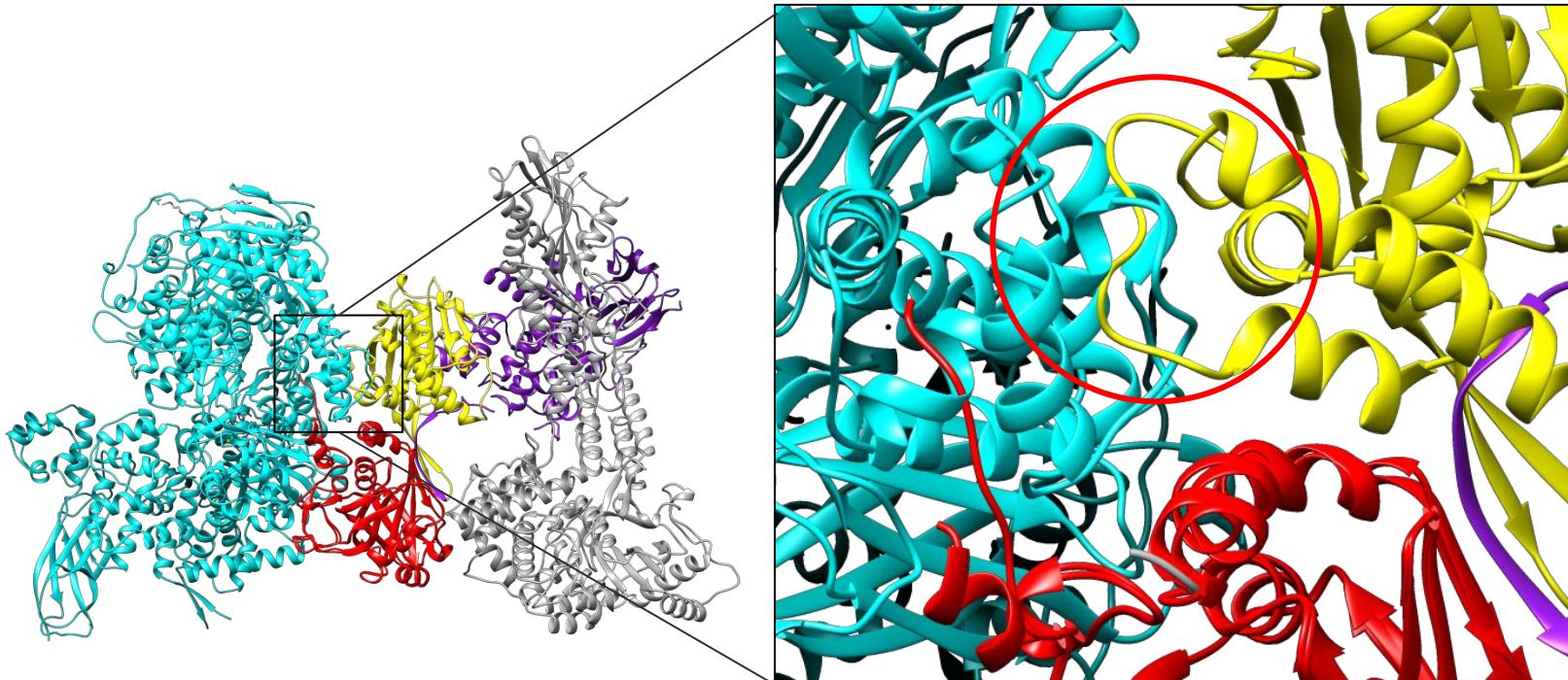


Transition Aar2-Brr2



PDB code: 4I43 and 4bdg
Resolution: 2A and 2.38A

Transition Aar2-Brr2

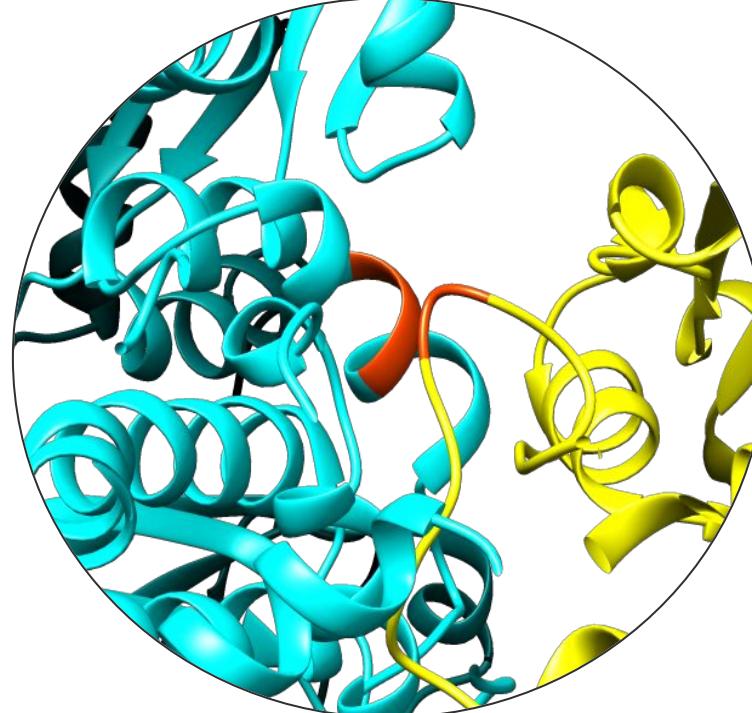
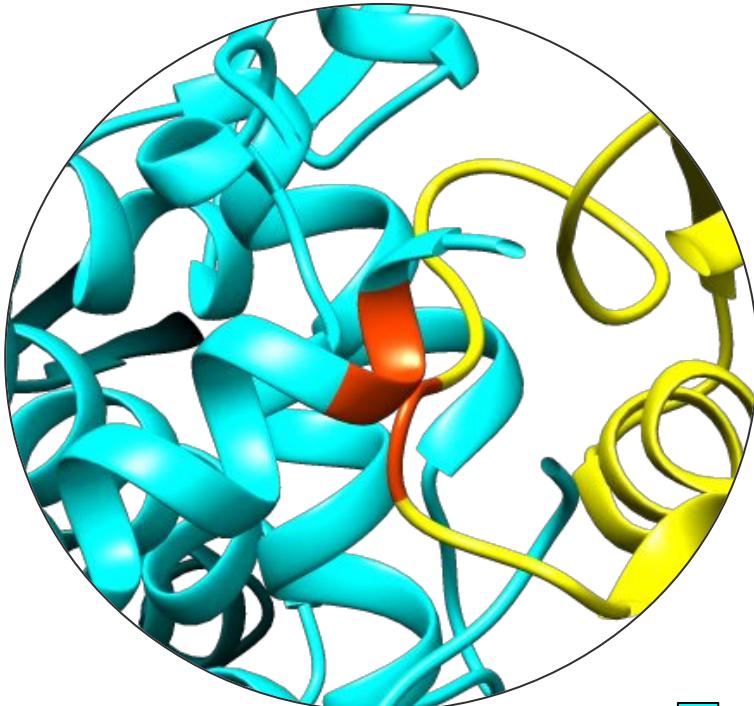


PDB code: 4I43 and 4bdg
Resolution: 2A and 2.38A

 Aar2
 Jab1/MPN

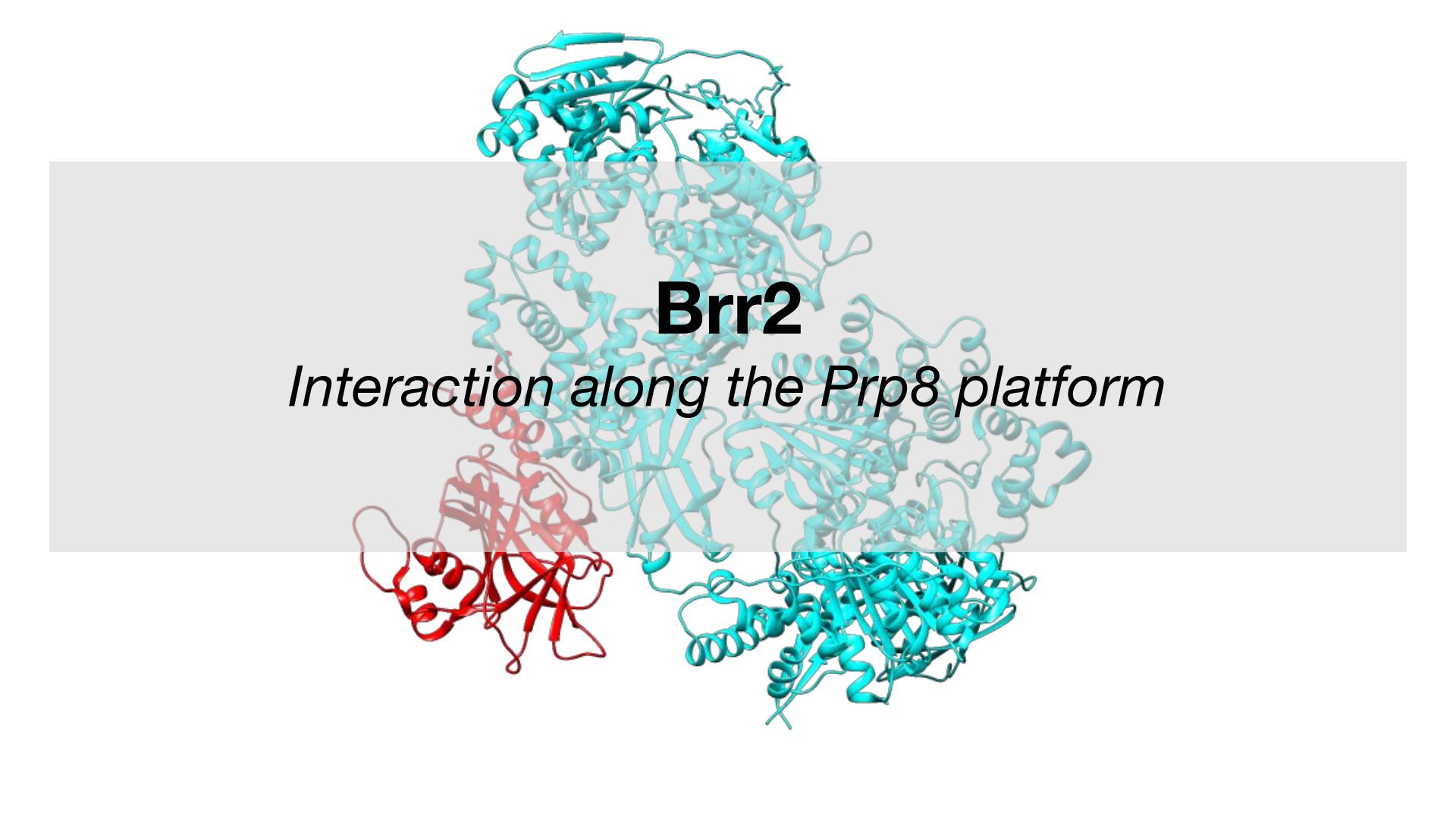
 Brr2
 RnaseH-like

Transition Aar2-Brr2



Brr2
 RnaseH-like

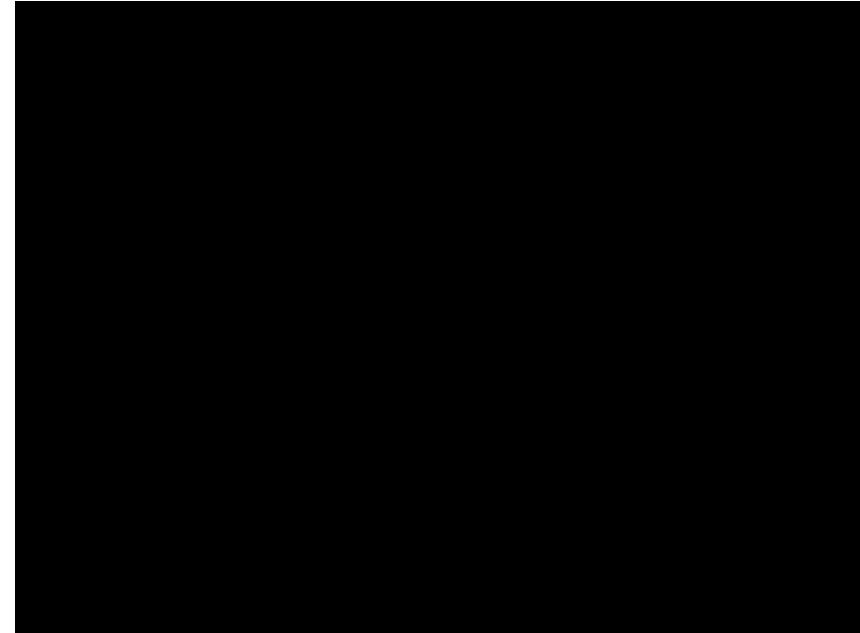
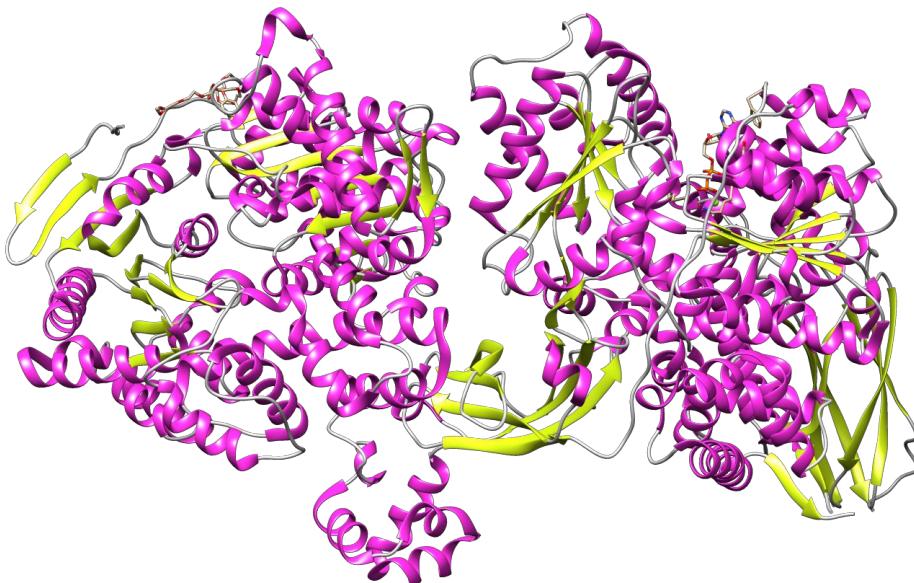
PDB code: 4I43 and 4bdg
Resolution: 2A and 2,38A



Brr2

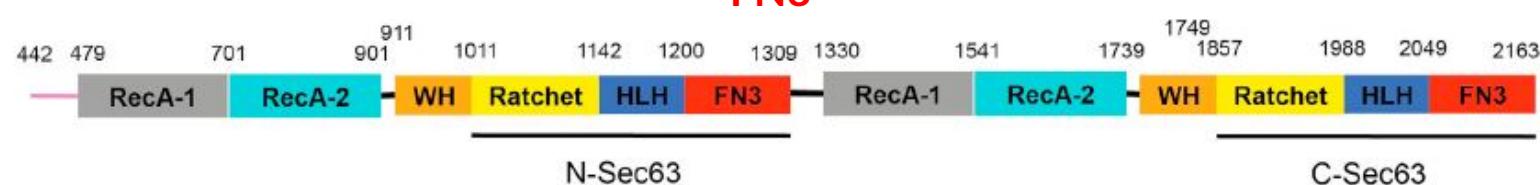
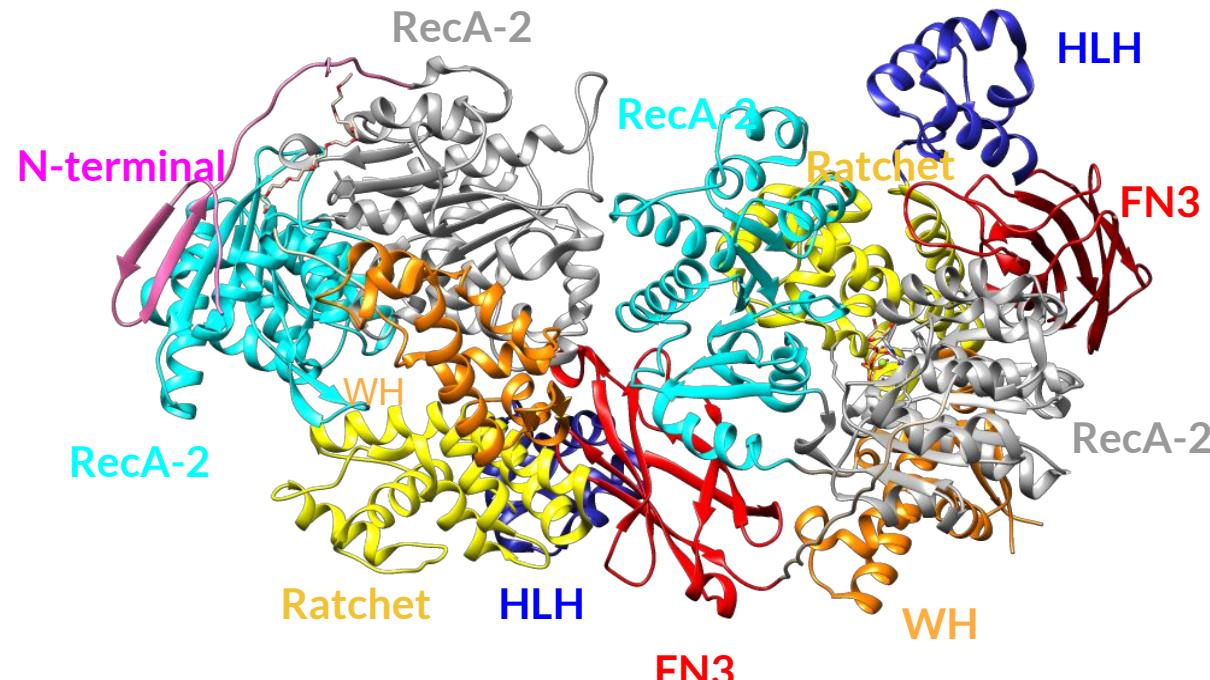
Interaction along the Prp8 platform

Brr2

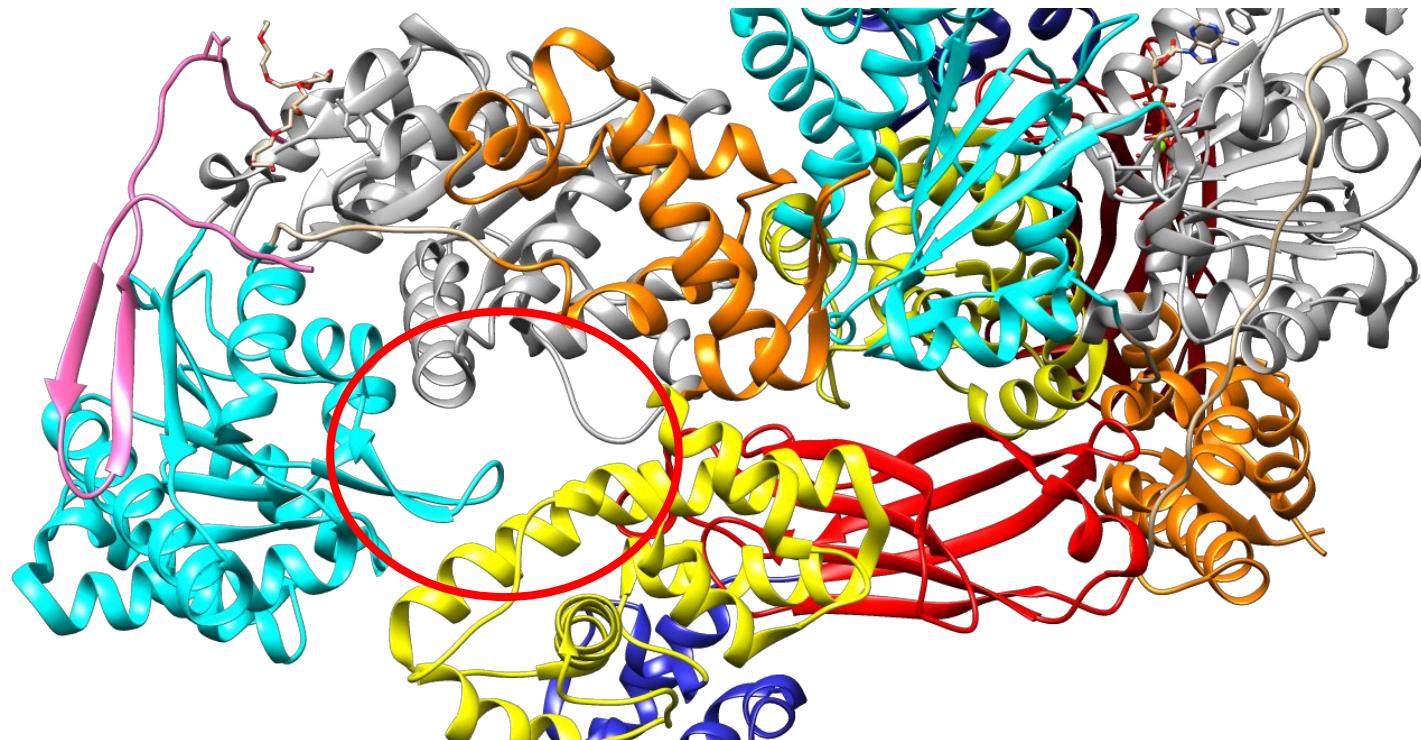


PDB code: 4BGD
Resolution: 2.84 Å

Brr2



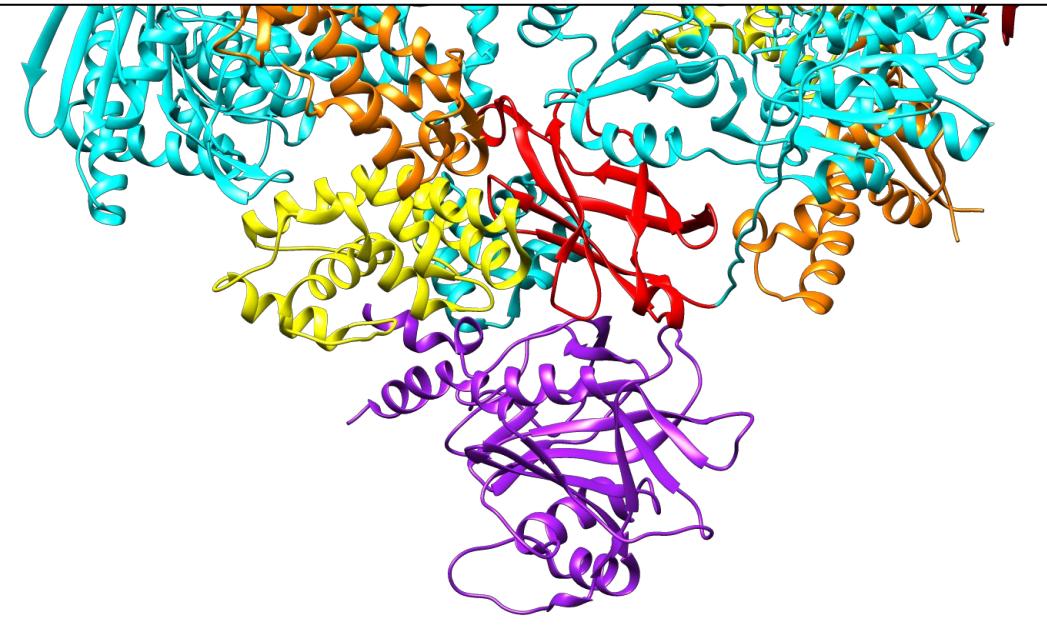
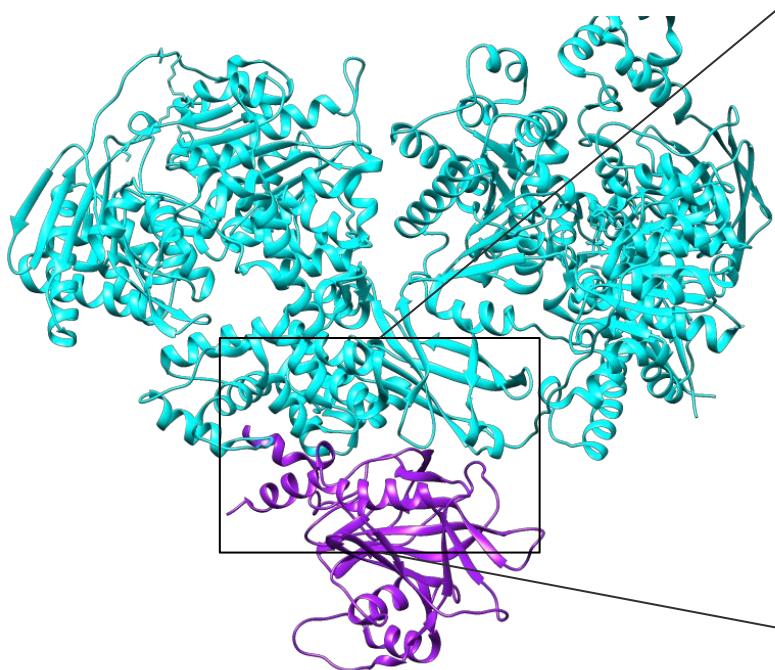
Brr2



RNA passage

PDB code: 4bgd
Resolution: 2.84A

Brr2

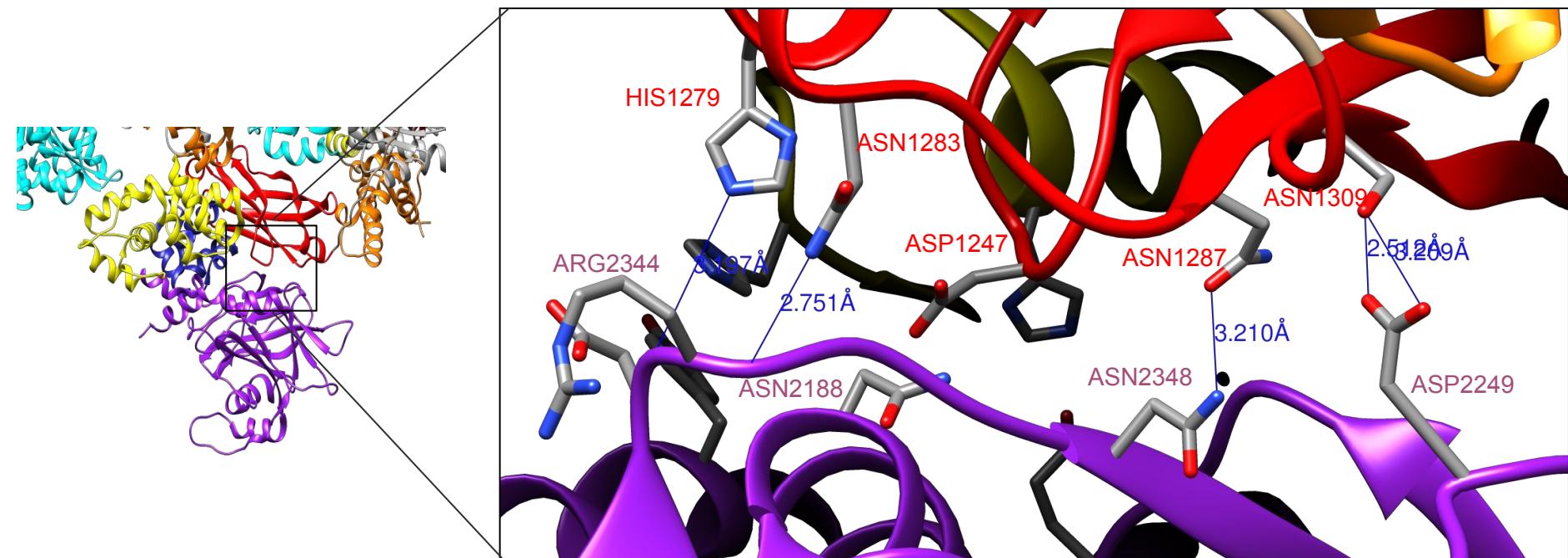


PDB code: 4bgd
Resolution: 2.84A

█	Brr2	█	FN3
█	Ratchet	█	Jab1/MPN

Brr2

Interaction Between Prp8 and the FN3 domain



Brr2

Asp 2343

Arg 2344

Jab1/MPN region

2341	2400				
LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFALL
LSDRFLGFFM	VPAQSSWNYN	FMGK.....
LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFALL
LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFALL
LSDRFLGFFM	VPAQGSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFALL
LSDRFLGFFM	VPGQVSWNYN	FMGVRHD..P	NMKYDLQLAN	PKEFYHEVHR	PSHFLNFASL
LSNKFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHELHR	TSHFLLFSNL
LSDRFLGYFM	VPSNGVWNYN	FQGQRWS..P	AMKFDVCLSN	PKEYYHEDHR	PVHFHNFKAF
LSDRITGNFI	IPSGNVWNYT	FMGTAFN..Q	EGDYNFKYGI	PLEFYNEMHR	PVHFLQFSEL
LSDRIQGFFL	VPEEGVWNYN	FNGASFS..P	KMTYSLKLDV	PLPFFALEHR	PTHVISYTEL
LSEKFRGFFL	VPDGGKWNYS	FMGSAGFg1E	KKPVHVVKLDT	PLPFYSDQHR	PIHFSSFNEL
LSDRFFGFYM	VPENGWPWNYN	FMGANHT..V	SINYSLTG	PKEYYHQVHR	PTHFLQFSKM
LSDRFLGFFM	VPAQASWNYN	FMGVRHD..P	NMKYELOLAN	PKEFYHEVHR	PSHFLNFALL

Negatively charged

Non-polar

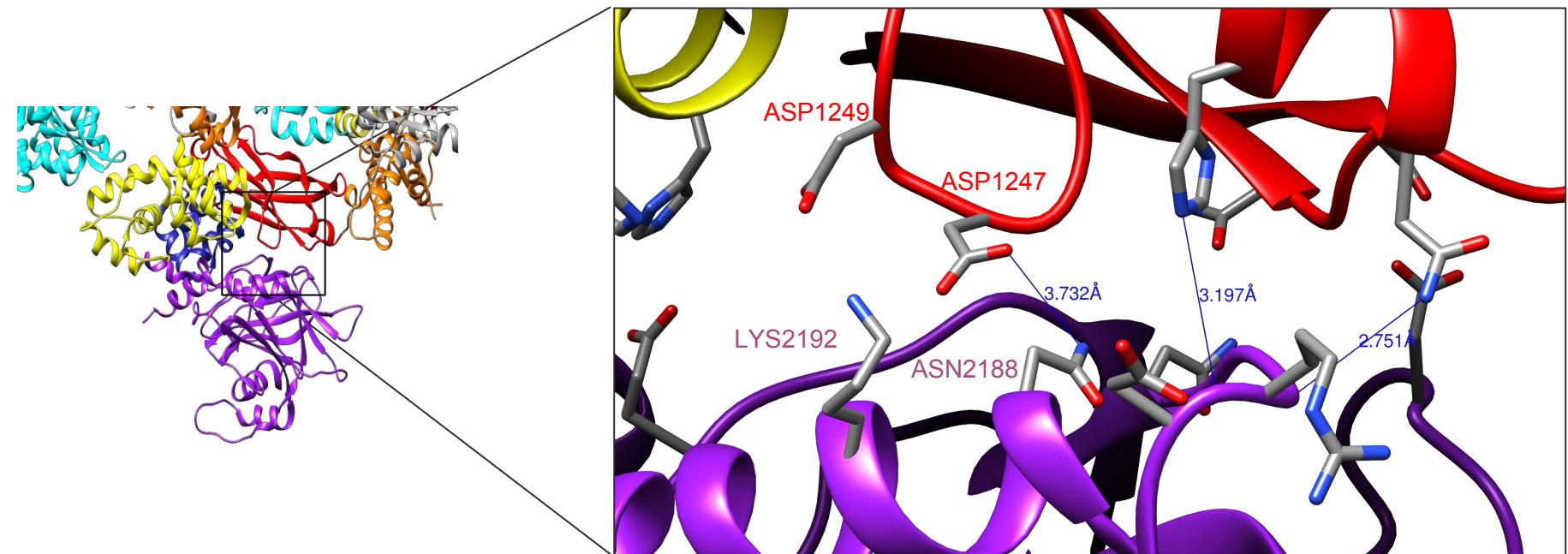
Positively charged

Polar

Sequence alignment based on PFAM domains

Brr2

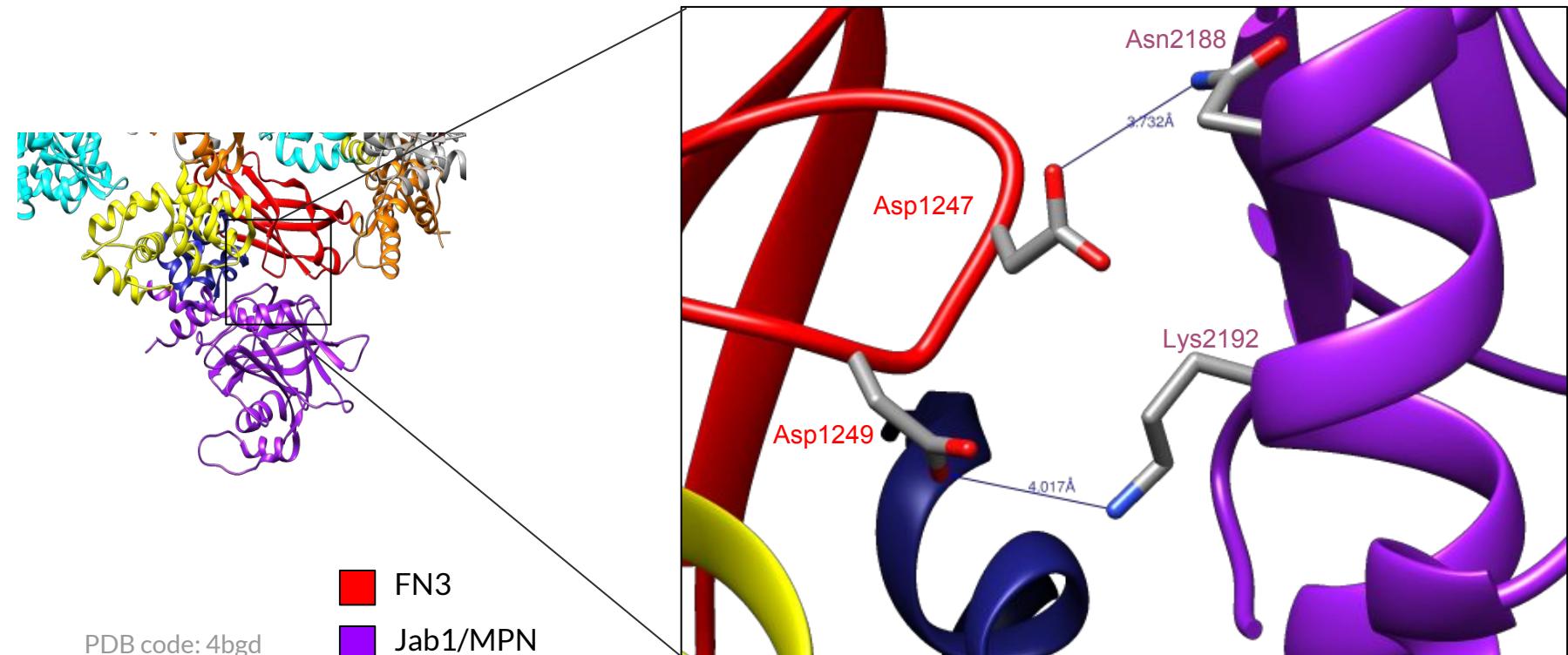
Interaction Between Prp8 and the FN3 domain



PDB code: 4bgd
Resolution: 2.84 Å

Brr2

Interaction Between Prp8 and the FN3 domain



Brr2

Jab1/MPN region

	2161	Asn 2188	Lys 2192	2220
H. sapiens	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
C. lupus	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
M. musculus	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
M. mulatta	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
G. gallus	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
X. tropicalis	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
D. rerio	ITSTTSNYET QTFSSKTEWR VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK			
D. melanogaster	ITSTTSNYET QTFSSKTEWR VRAISATNLH LRTNHIYVSS DDIKETG.YT YILPKNILKK			
C. elegans	ITATTTSNYET ASFASRTEWR VRAISSTNLH LRTQHIYVNS DDVKDTG.YT YILPKNILKK			
A. thaliana	ISTTISPYEQ SAFGSKTDWR VRAISATNLY LRVNHIYVNS DDIKETG.YT YIMPKNILKK			
S. pombe	VVTTTSAYEN EKFSSKTEWR NRAISSISLP LRTKNIYVNS DNISETFPYT YILPQNLLRK			
N. crassa	IVTTTSQFEQ QTFASKTEWR TRAIATSNLR TRANNMYVSP VDSLDD.VT YVMPKNILKR			
S. cerevisiae	VVVASADYES QTFSSKNEWR KSAIANTLLY LRLKNIYVSA DDFVEEQ.NV YVLPKNLLKK			

Negatively charged

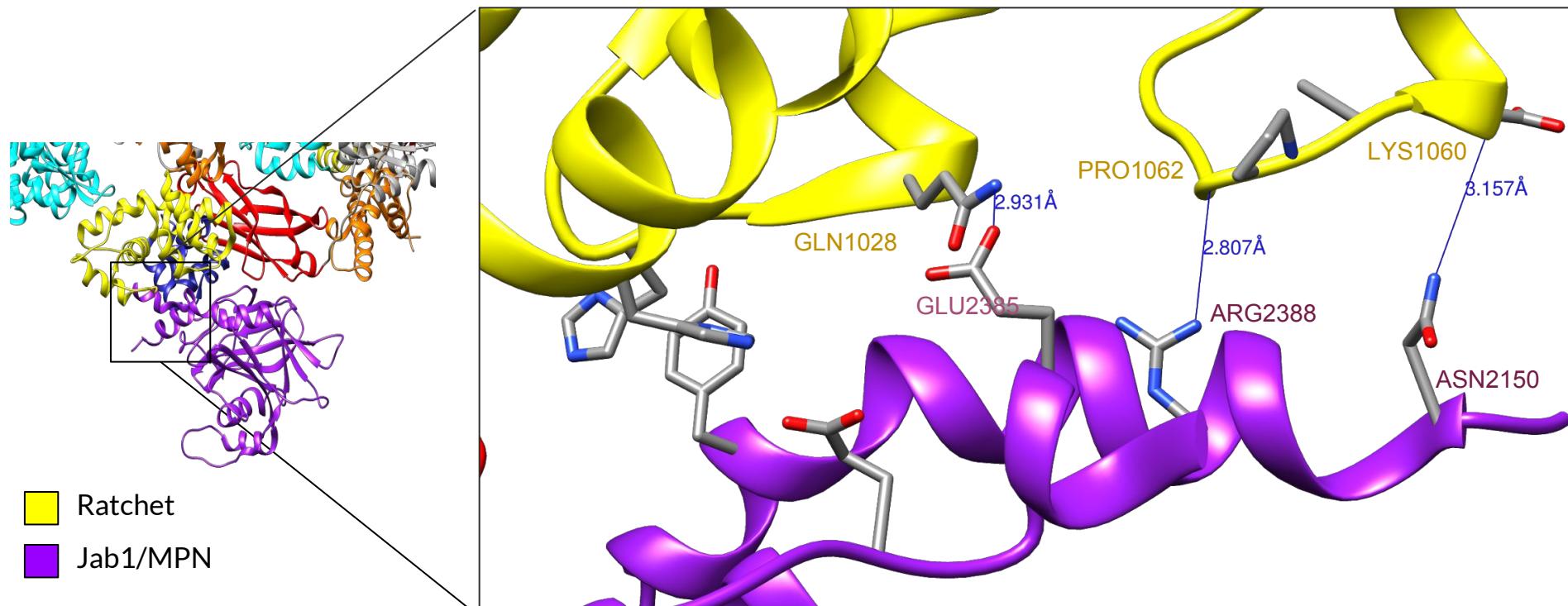
Non-polar

Positively charged

Polar

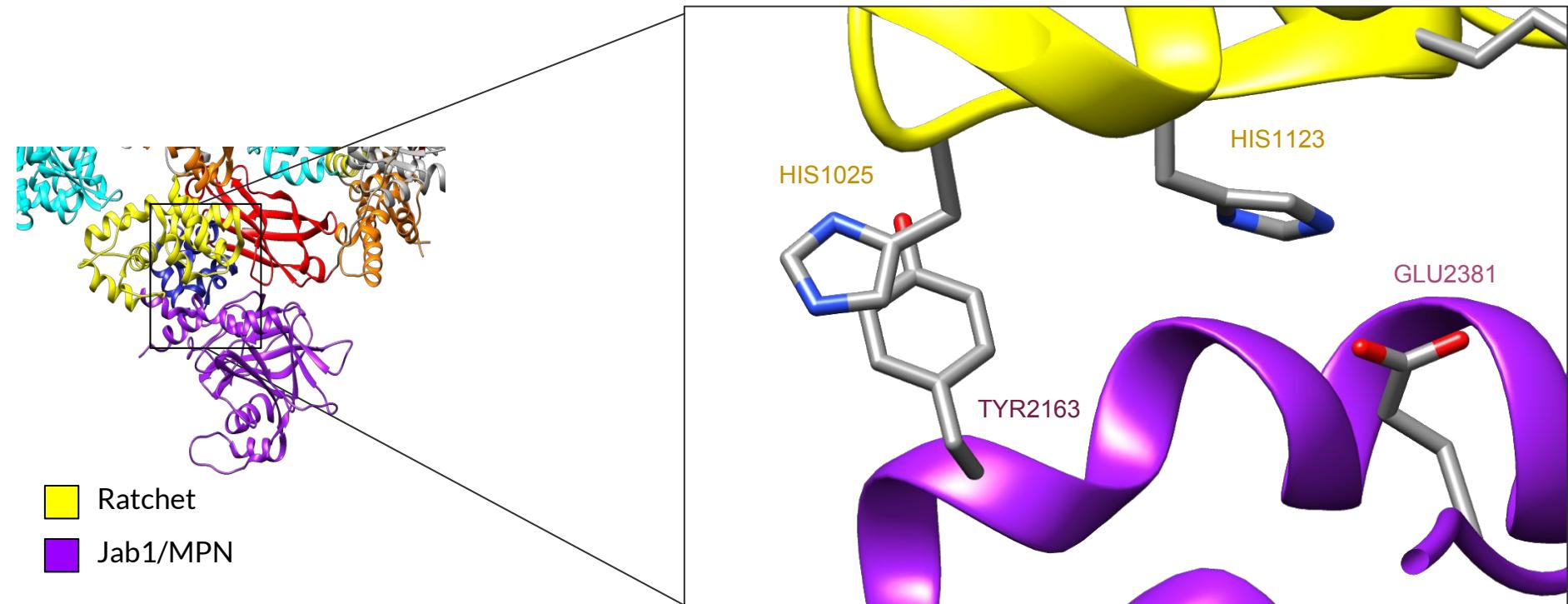
Clustal Sequence alignment

Brr2



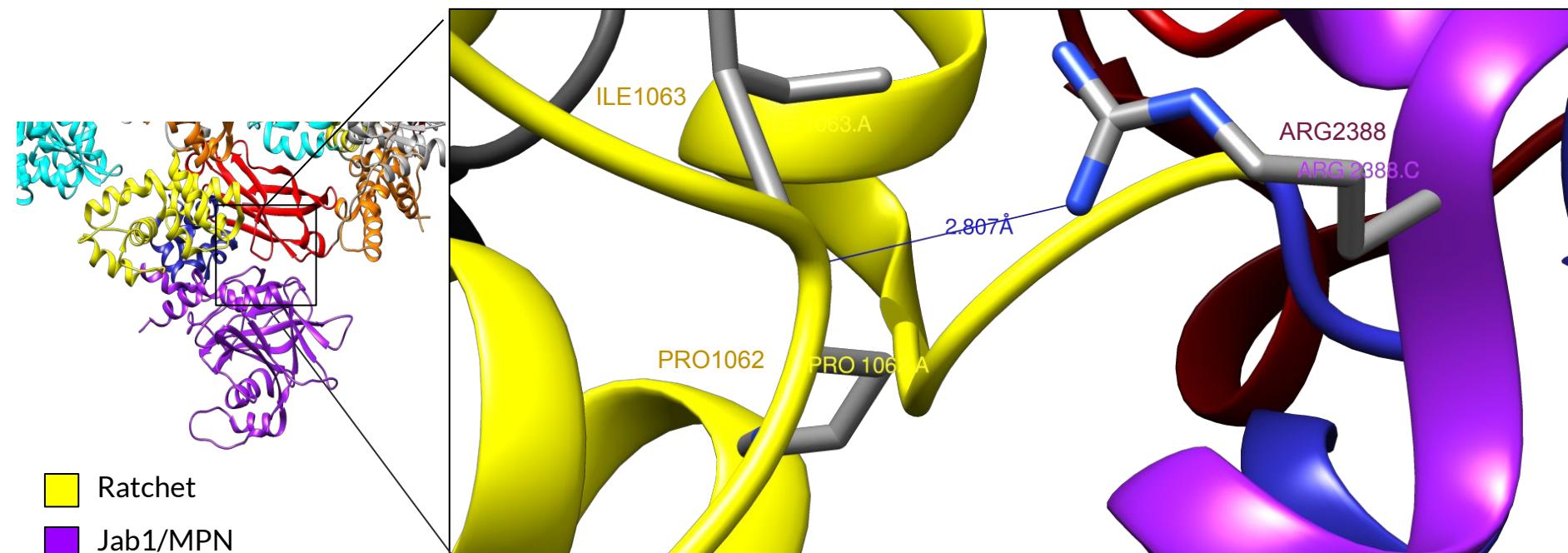
PDB code: 4bgd
Resolution: 2.84 Å

Brr2



PDB code: 4bgd
Resolution: 2.84A

Brr2



PDB code: 4bgd
Resolution: 2.84 Å

Brr2

		<u>Jab1/MPN region</u>				Glu 2381 Glu 2385	Arg 2388	
	2341						*	2400
H. sapiens	LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR		PSHFLNFALL	
M. mulatta	LSDRFLGFFM	VPAQSSWNYN	FMGK.....
C. lupus	LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR		PSHFLNFALL	
M. musculus	LSDRFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR		PSHFLNFALL	
G. gallus	LSDRFLGFFM	VPAQGSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR		PSHFLNFALL	
D. rerio	LSDRFLGFFM	VPGQVSWNYN	FMGVRHD..P	NMKYDLQLAN	PKEFYHEVHR		PSHFLNFAVL	
D. melanogaster	LSNKFLGFFM	VPAQSSWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHELHR		TSHFLLFSNL	
C. elegans	LSDRFLGYFM	VPSNGVWNYN	FQGQRWS..P	AMKFDVCLSN	PKEYYHEDHR		PVHFHNFKAF	
S. cerevisiae	LSDRITGNFI	IPSGNVWNYT	FMGTAFN..Q	EGDYNFKYGI	PLEFYNEMHR		PVHFLQFSEL	
S. pombe	LSDRIQGFFL	VPEEGVWNYN	FNGASFS..P	KMTYSLKLDV	PLPFFALEHR		PTHVISYTEL	
N. crassa	LSEKFRGFLL	VPDGGKWNYS	FMGSAGFgle	KKPVHVVKLDT	PLPFYSDQHR		PIHFSSFNEL	
A. thaliana	LSDRFFGFYM	VPENGPNWYN	FMGANHT..V	SINYSLTLGT	PKEYYHQVHR		PTHFLQFSKM	
X. tropicalis	LSDRFLGFFM	VPAQASWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR		PSHFLNFALL	

Negatively charged

Non-polar

*RP13

Positively charged

Polar

Sequence alignment based on PFAM domains

Brr2

Asn 2150

Jab1/MPN region

	2161	Asn 2150	2220
H.sapiens	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
C.lupus	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
M.musculus	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
M.mulatta	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
G.gallus	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
X.tropicalis	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
D.rerio	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH LRTNHIYVSS DDIKETG.YT YILPKNVLK
D.melanogaster	ITSTTSNYET	QTFSSKTEWR	VRAISATNLH LRTNHIYVSS DDIKETG.YT YILPKNVLKK
C.elegans	ITATTNSNYET	ASFASRTEWR	VRAISSTNLH LRTQHIYVNS DDVKDTG.YT YILPKNVLKK
A.thaliana	ISTTISPYEQ	SAFGSKTDWR	VRAISATNLY LRVNHIYVNS DDIKETG.YT YIMPKNILKK
S.pombe	VVTTTSAYEN	EKFSSKTEWR	NRAISSLISLP LRTKNIYVNS DNISETFPYT YILPQNLLRK
N.crassa	IVTTTSQFEQ	QTFASKTEWR	TRAIATSNLR TRANNMVYVSP VDSLDD.VT YVMPKNILKR
S.cerevisiae	VVVASADYES	QTFSSKNEWR	KSAIANTLLY LRLKNIYVSA DDFVEEQ.NV YVLPKNLLKK

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

Conclusions

- Prp8 represents a **fundamental platform** for the splicing reactions. Its conservation through species is the first prove of its important role in cells.
- Nevertheless, further studies are needed to fully understand the **whole spliceosomal protein complex**.
- Although its complicated nature makes it proper difficult to understand.
- The **transition from Aar2 to Brr2** is crucial in order to finally form the catalytic complex.
- Interaction of Brr2 with Prp8 plays a key role in the formation of the **catalytic centre of the spliceosome**.

Bibliography

- ❑ Galej WP, Nguyen T, J Newman, Nagai K. Structural studies of the spliceosome: zooming into the heart of the machine. *Curr Opin Struct Biol.* 2014 Apr; 25(100): 57–66.
- ❑ Galej, W. P., Oubridge, C., Newman, A. J. & Nagai, K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature* 493, 638–643 (2013)
- ❑ Nguyen TH, Galej WP, Bai X, et al. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature*. 2016. (7590):298-302.
- ❑ Shi Y. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nature Reviews Molecular Cell Biology*. 2017;18, 655-670.
- ❑ Wahl M, Will CL, Lührmann. The Spliceosome: Design Principles of Dynamic RNP Machine. 2009. *Cell*. 136, 701-718.
- ❑ Will CL, Lührmann R. Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology*. 2011;3(7):a003707. doi:10.1101/cshperspect.a003707.
- ❑ Yan, C. et al. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* 349, 1182–1191 (2015)
- ❑ MN Luscombe, Laskowski AR, Thornton MJ. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Research*. 2001 (29); 2860-2874.
- ❑ Nguyen, Thi Hoang Duong, et al. Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site. *Structure* 2013 (21.6); 910-919.

Questions

1. The proteins Aar2 and Brr2 of the spliceosome:
 - a) Are mutually exclusive
 - b) Interact in the spliceosome
 - c) Doesn't exist
 - d) Are not important
 - e) All the others
2. Regarding Brr2 protein in the spliceosome, which of the following is wrong:
 - a) It only interacts with the Jab1/MPN domain
 - b) Brr2 is disassembled when Prp8 gets inside the nucleus.
 - c) It's the only ski2-like helicase in the spliceosome
 - d) It is a very large protein
 - e) Aar2 is not longer needed once the Prp8 complex goes inside the nucleus.
3. Regarding Prp8:
 - a) It's one of the most conserved proteins in the spliceosome
 - b) It is not important for the spliceosome
 - c) The last two ones
 - d) It only has one domain
 - e) All the others

4. Which one of the following domains is present in *S.cerevisiae* Prp8?

- a. N-terminal
- b. RNase-H like
- c. Thumb/X
- d. Jab1/MPN
- e. Tots els anteriors.

5. What type of residues interact better with RNA?

- a. Non-polar
- b. Negatively charged
- c. A and C
- d. Positively charged
- e. RNA doesn't interact with proteins.

6. Which of the following proteins contains a GTP ligand?

- 1. Snu114
- 2. Brr2
- 3. Aar2
- 4. U1 snRNA

7. Which of the following snRNP is the first to recognise the 5'-splice site (5'-SS)?

1. U1
2. U2
3. U3
4. U4

8. In U5 snRNA recognition, which of the following Prp8 domains is crucial?

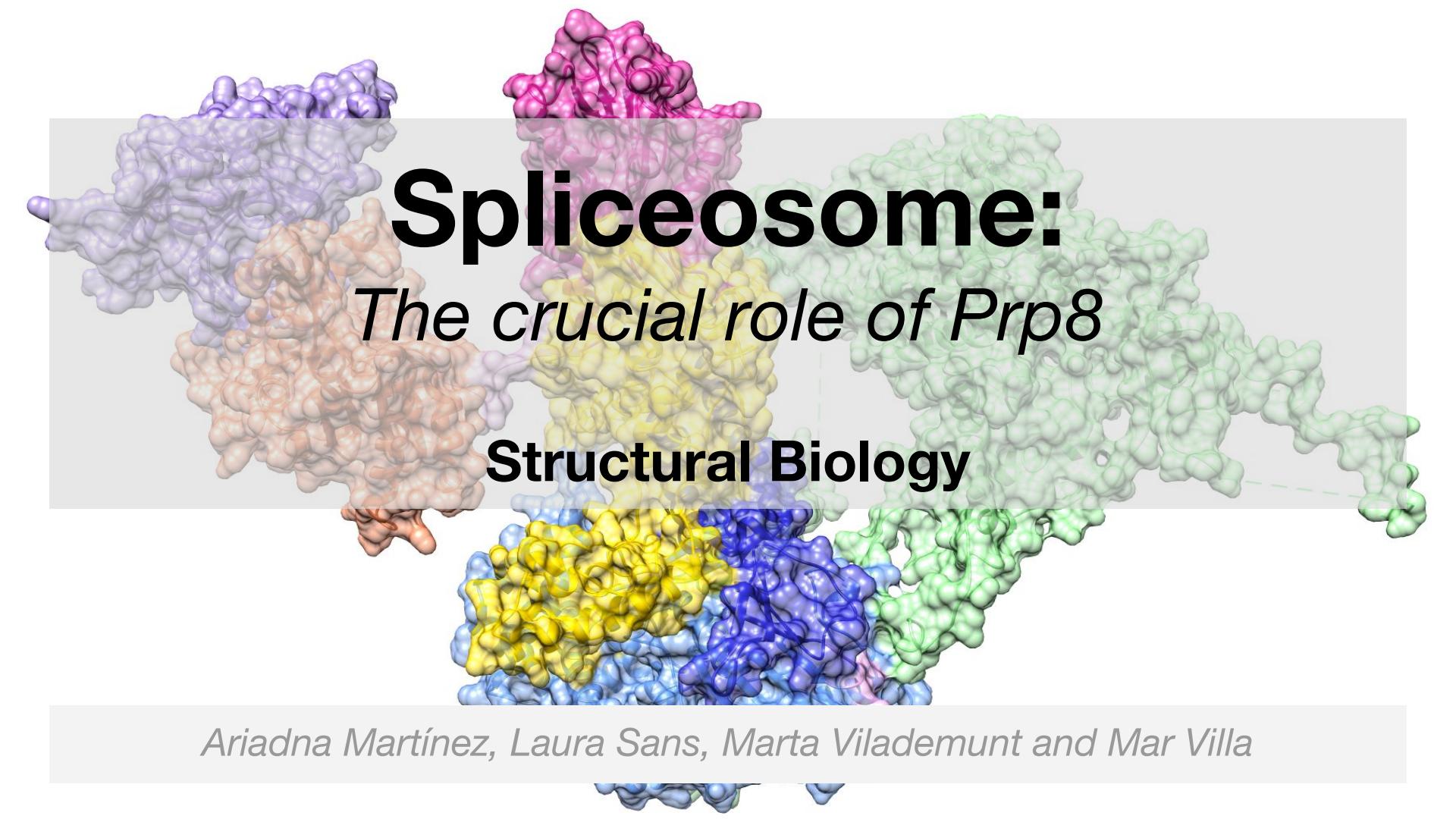
1. N-terminal domain
2. Jab1/MPN
3. RNase H like
4. None of them

9. Regarding the RH domain of Prp8, which of the following is true:

- a. Suffers a conformational change so it offers a Mg²⁺⁺ ion to the splicing active site.
- b. In the close conformation the RH domain forms a loose loop.
- c. Is in the close conformation when the RH domain can host a Mg²⁺⁺ ion.
- d. RH doesn't suffer any conformational change at all, that's why it's called RH-like.
- e. None of them is true.

10. Which part of Aar2 allows the different domains of Prp8 to interact?

- a) The tail
- b) N-terminal
- c) Betta domain
- d) Zing domain
- e) The loops



Spliceosome:

The crucial role of Prp8

Structural Biology

Ariadna Martínez, Laura Sans, Marta Vilademunt and Mar Villa

STAMP Structural Alignment of Multiple Proteins

Version 4.4 (May 2010)

by Robert B. Russell & Geoffrey J. Barton

Please cite PROTEINS, v14, 309-323, 1992

Running roughfit.

Sc = STAMP score, RMS = RMS deviation, Align = alignment length

Len1, Len2 = length of domain, Nfit = residues fitted

Secs = no. equivalent sec. strucs. Eq = no. equivalent residues

%I = seq. identity, %S = sec. str. identity

P(m) = P value (p=1/10) calculated after Murzin (1993), JMB, 230, 689-694

(NC = P value not calculated - potential FP overflow)

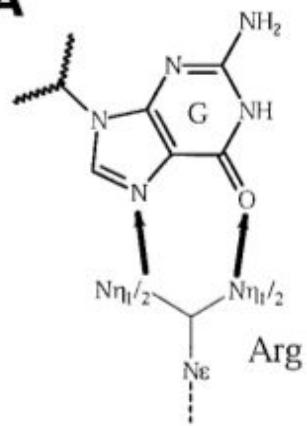
No.	Domain1	Domain2	Sc	RMS	Len1	Len2	Align	NFit	Eq.	Secs.	%I	%S	P(m)
Pair 1	RNAopen	RNAclose	7.40	0.93	221	221	239	188	186	0	98.39	100.00	0.00e+00

Reading in matrix file RNAse.mat...

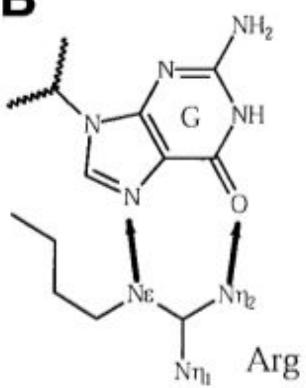
Doing cluster analysis...

Cluster: 1 (RNAopen & RNAclose) Sc 7.39 RMS 0.93 Len 239 nfit 188

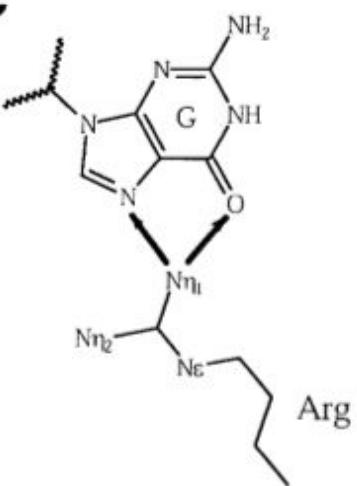
See file RNAse.1 for the alignment and transformations

A

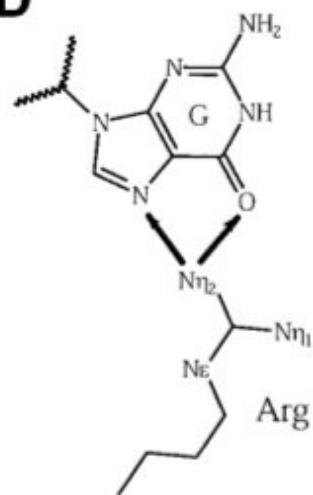
(21 examples, 47 bonds)

B

(2 examples, 4 bonds)

C

(2 examples, 4 bonds)

D

(4 examples, 8 bonds)

Prp8

STAMP Structural Alignment of Multiple Proteins

Version 4.4 (May 2010)
by Robert B. Russell & Geoffrey J. Barton
Please cite PROTEINS, v14, 309-323, 1992

Sc = STAMP score, RMS = RMS deviation, Align = alignment length
Len1, Len2 = length of domain, Nfit = residues fitted
Sects = no. equivalent sec. strucs. Eq = no. equivalent residues
%I = seq. identity, %S = sec. str. identity
P(m) = P value (p=1/10) calculated after Murzin (1993), JMB, 230, 689-694
(NC = P value not calculated - potential FP overflow)

Cluster: 1 (hENDO & yENDO) Sc 9.12 RMS 1.19 Len 172 nfit 172
See file templates_OK.1 for the alignment and transformations
Cluster: 2 (hLINKER & yLINKER) Sc 8.43 RMS 1.42 Len 278 nfit 264
See file templates_OK.2 for the alignment and transformations
Cluster: 3 (hJAB & yJAB) Sc 8.02 RMS 1.29 Len 248 nfit 228
See file templates_OK.3 for the alignment and transformations
Cluster: 4 (hRNase & yRNase) Sc 7.11 RMS 1.24 Len 192 nfit 190
See file templates_OK.4 for the alignment and transformations
Cluster: 5 (hRT & yRT) Sc 6.94 RMS 0.90 Len 367 nfit 364
See file templates_OK.5 for the alignment and transformations
Cluster: 6 (hNT & yNT) Sc 6.53 RMS 1.58 Len 779 nfit 590
See file templates_OK.6 for the alignment and transformations
Cluster: 7 (hENDO yENDO & hJAB yJAB) Sc 1.45 RMS 3.88 Len 267 nfit 56 LOW SCORE
See file templates_OK.7 for the alignment and transformations
Cluster: 8 (hENDO yENDO hJAB yJAB & hRNase yRNase) Sc 0.92 RMS 4.69 Len 400 nfit 30 LOW SCORE
See file templates_OK.8 for the alignment and transformations
Cluster: 9 (hLINKER yLINKER & hENDO yENDO hJAB yJAB hRNase yRNase) Sc 1.04 RMS 5.04 Len 540 nfit 21 LOW SCORE
See file templates_OK.9 for the alignment and transformations
Cluster: 10 (hLINKER yLINKER hENDO yENDO hJAB yJAB hRNase yRNase & hRT yRT) Sc 0.62 RMS 4.69 Len 746 nfit 18 LOW SCORE
See file templates_OK.10 for the alignment and transformations
Cluster: 11 (hLINKER yLINKER hENDO yENDO hJAB yJAB hRNase yRNase hRT yRT & hNT yNT) Sc 0.63 RMS 5.31 Len 1303 nfit 10 LOW SCORE
See file templates OK.11 for the alignment and transformations