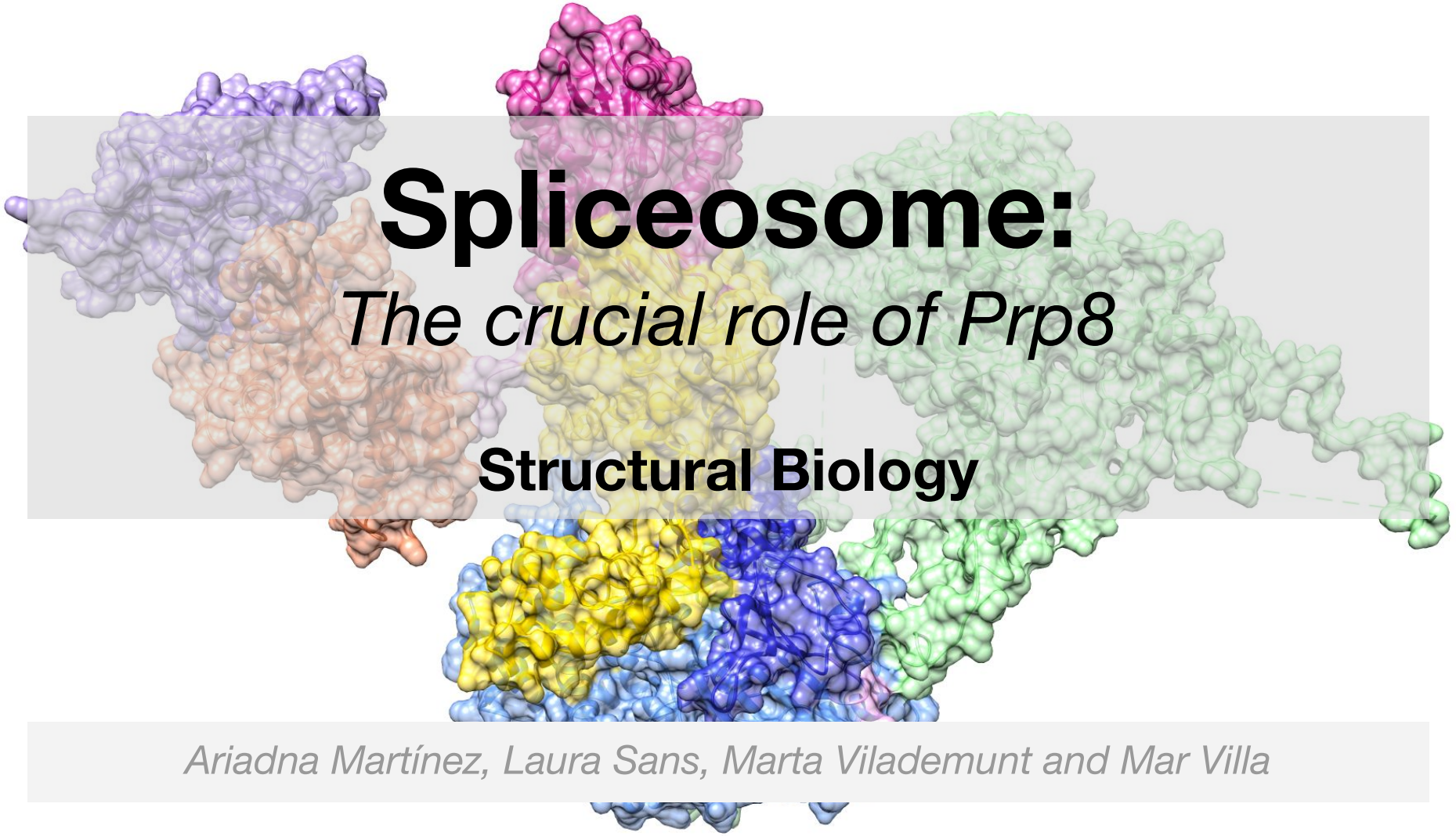


Spliceosome:

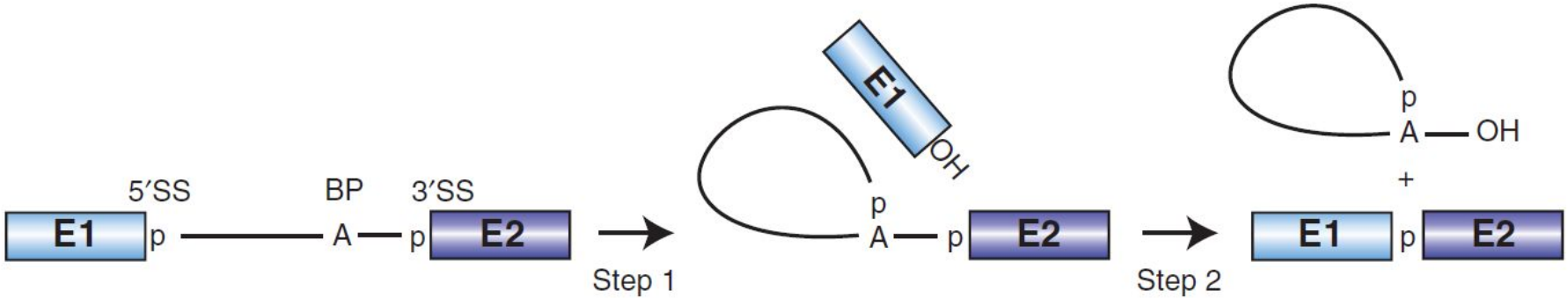
The crucial role of Prp8

Structural Biology

Ariadna Martínez, Laura Sans, Marta Vilademunt and Mar Villa

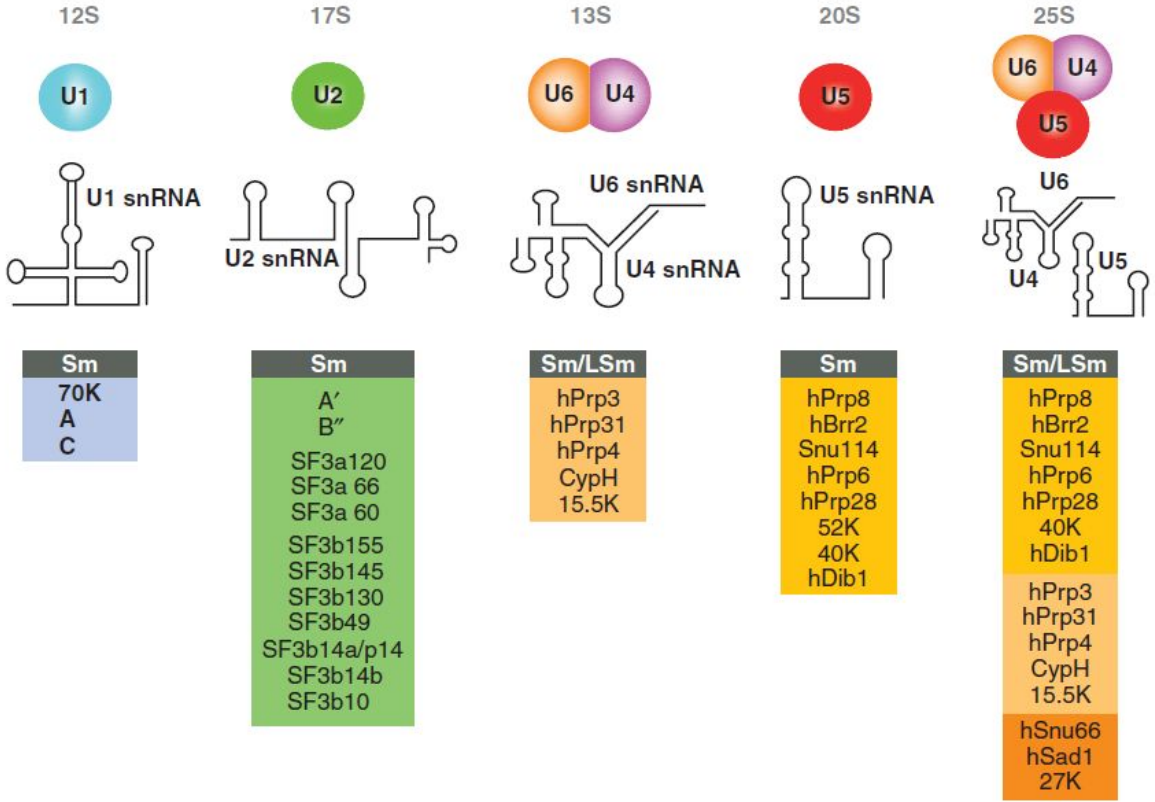
Index

Spliceosome basics - **Branching and exon ligation**



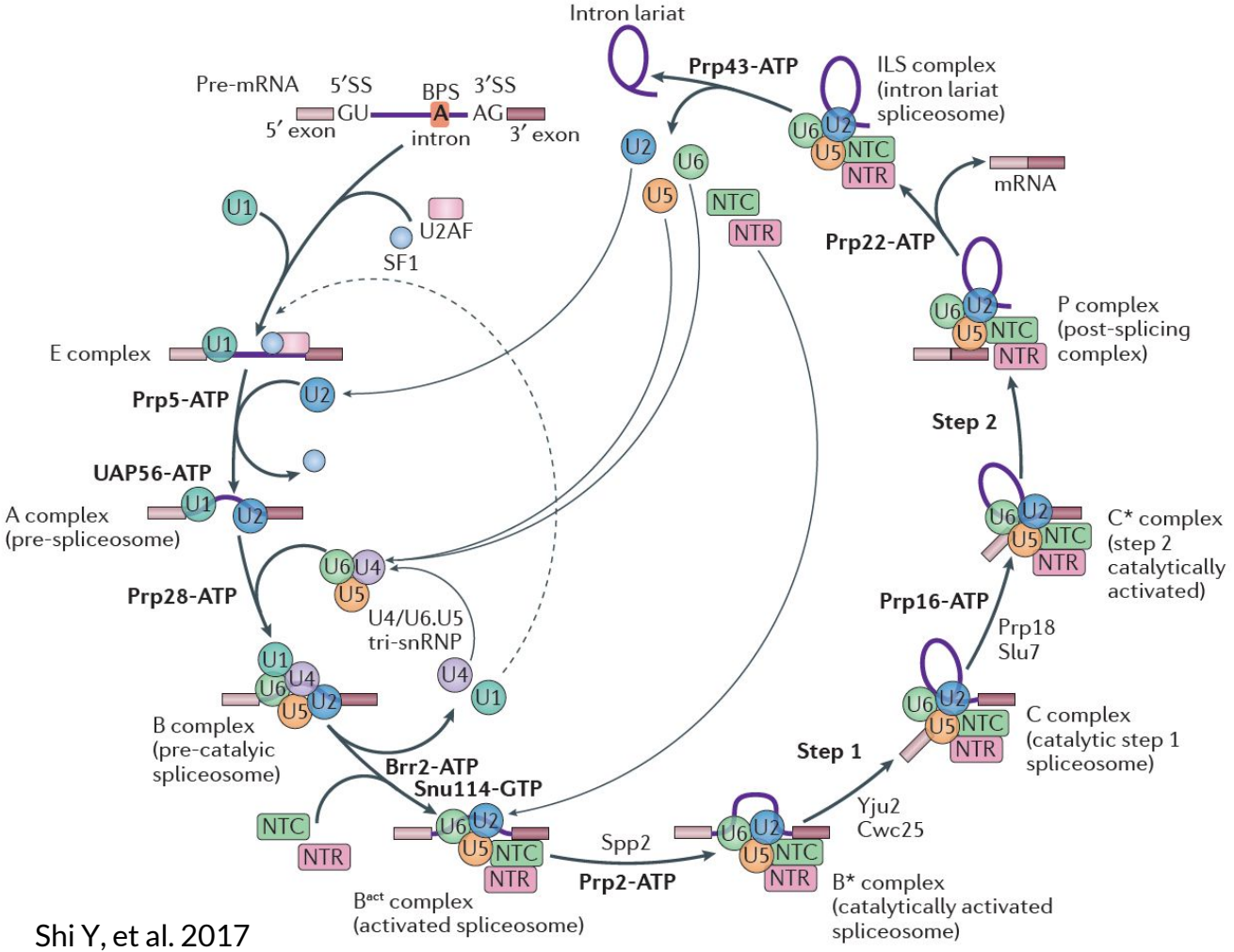
Will CL, et al. 2011

Spliceosome basics - **Branching and exon ligation**



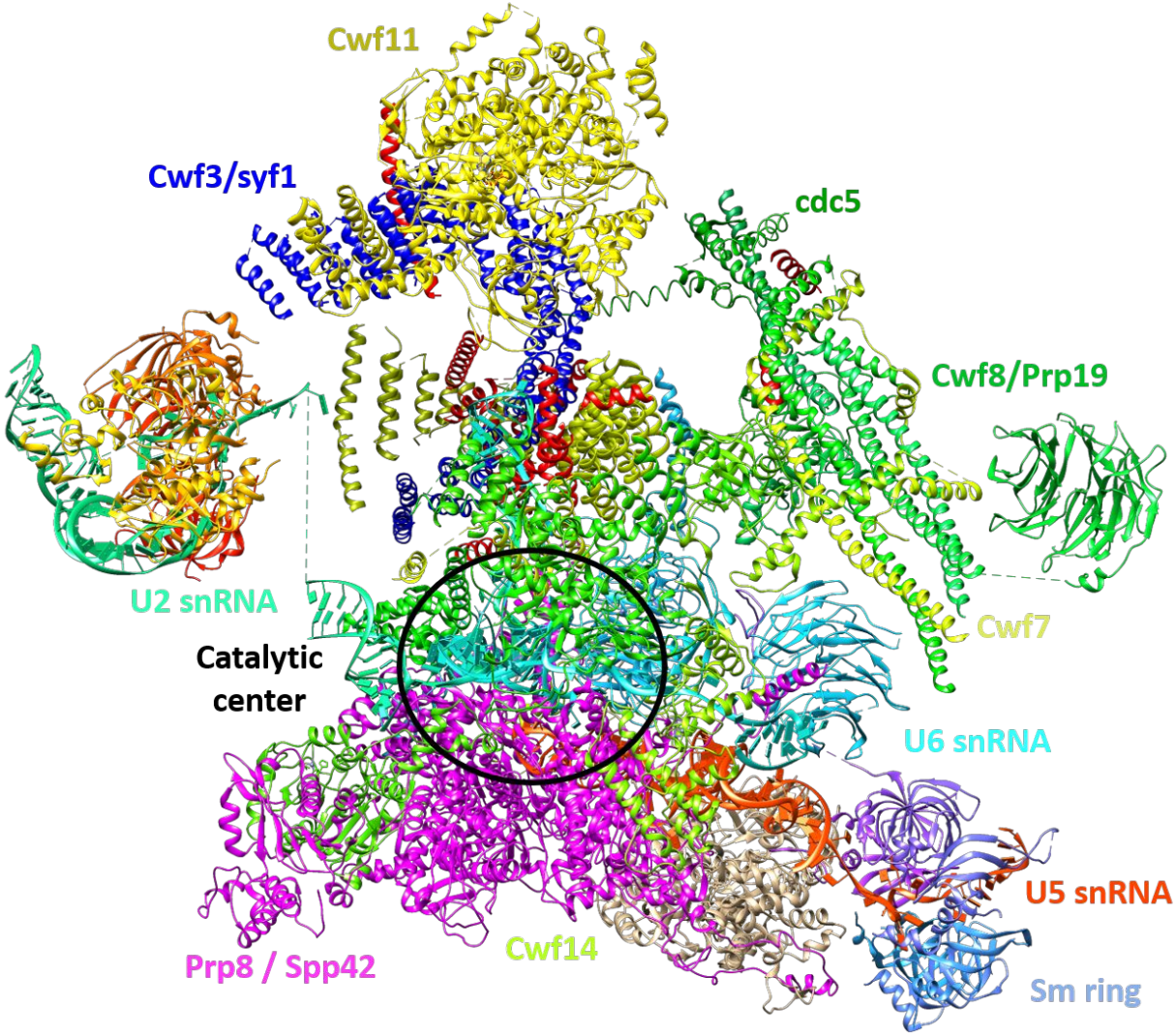
Will CL, et al. 2011

Spliceosome basics - **STEPS**



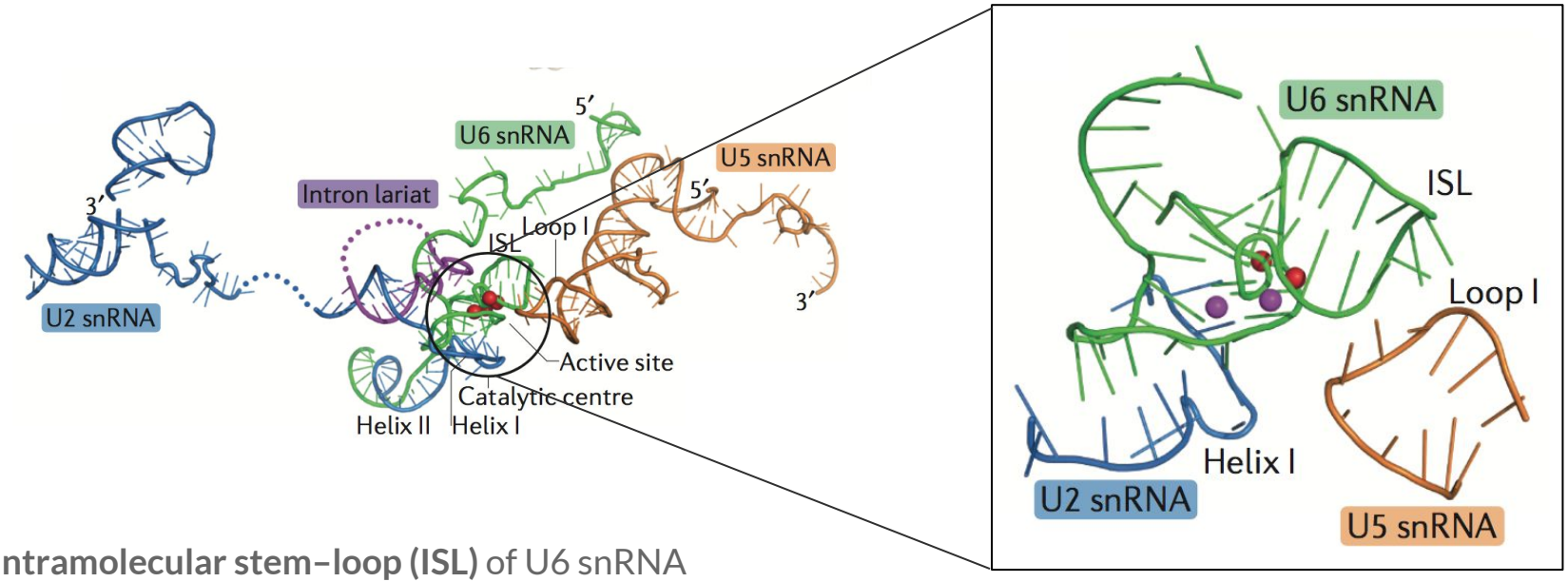
Shi Y, et al. 2017

Structure of yeast spliceosome



PDB code: 3JB9
Resolution: 3,6 A

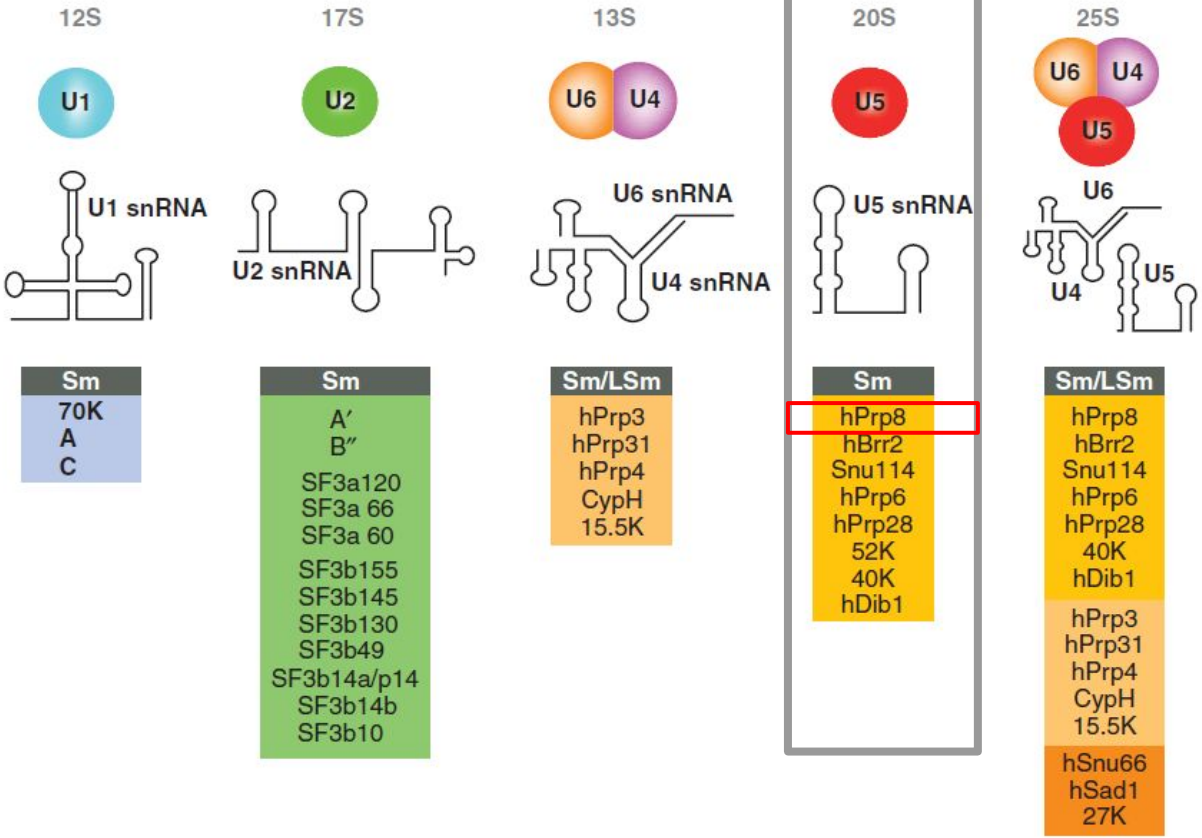
Catalytic core



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

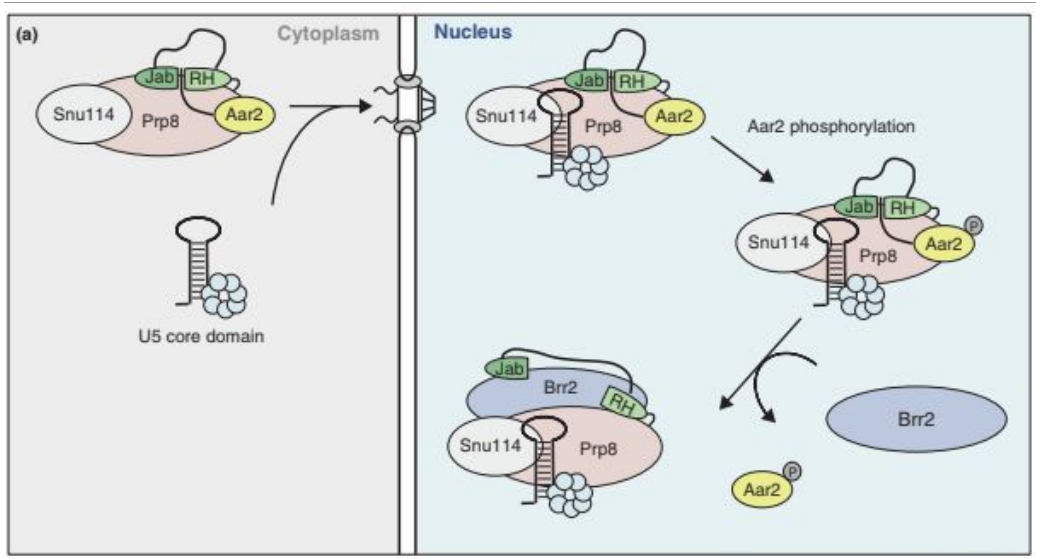
Shi Y, et al. 2017

U5 snRNP

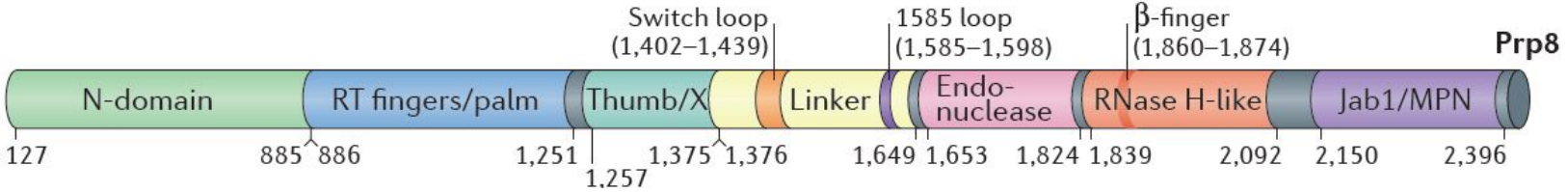


Will CL, et al. 2011

Prp8

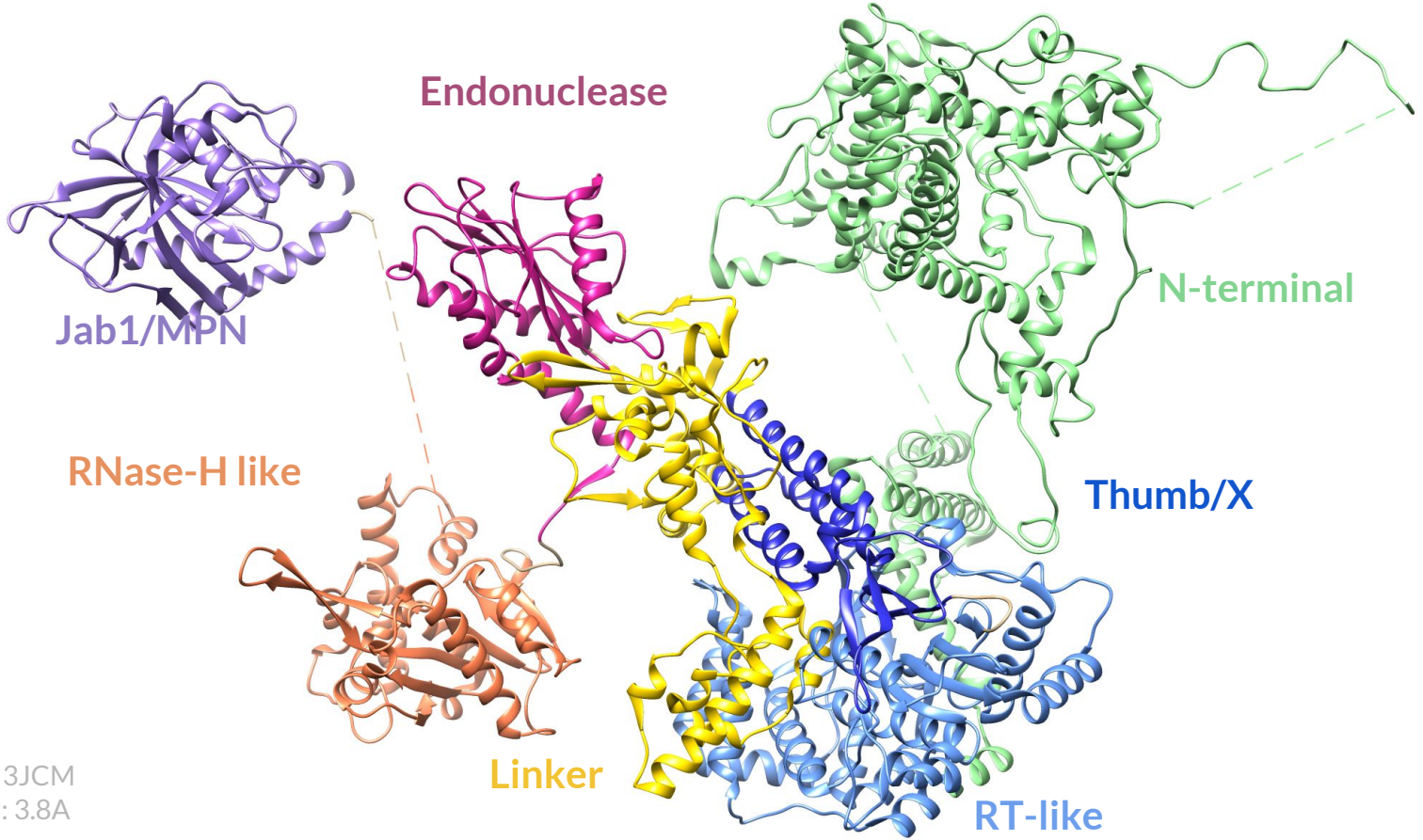


Galej, et al. 2014



Shi Y, et al. 2017

Prp8



PDB code: 3JCM
Resolution: 3.8A

Prp8

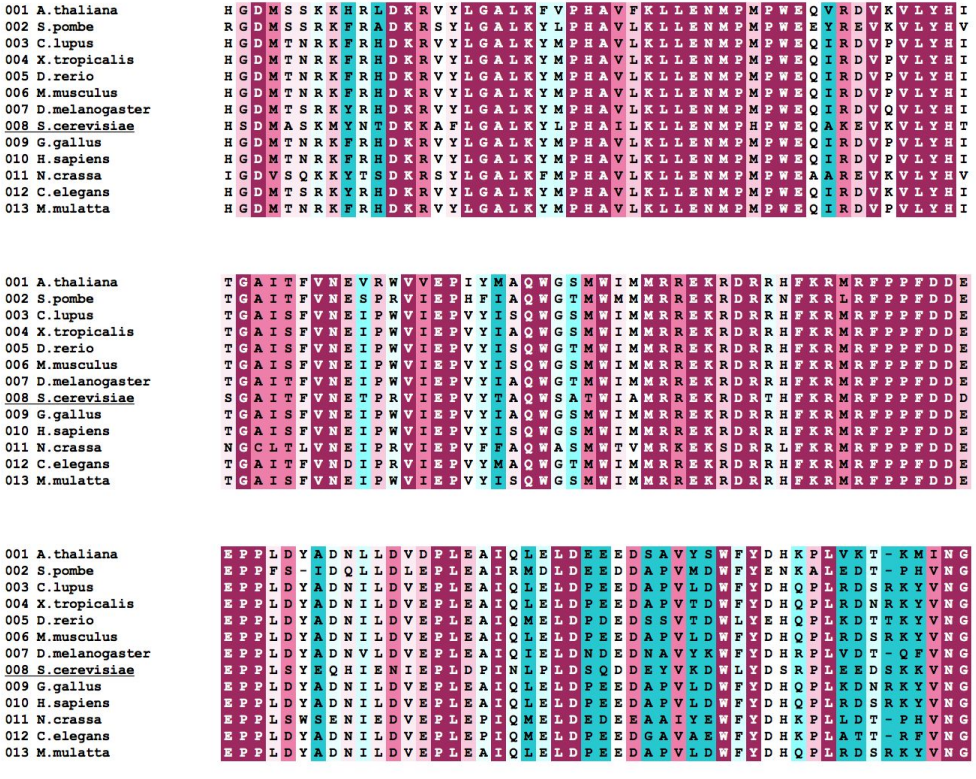
Class	Alpha and beta proteins
Fold	Ribonuclease H-like motif *(3 layers: a/b/a; mixed beta-sheet of 5 strands, order 32145; strand 2 is antiparallel to the rest)*
Superfamily	Ribonuclease H-like
Family	Prp8 beta-finger domain-like
Protein	Pre-mRNA-splicing factor 8
Species	*Saccharomyces cerevisiae*

Prp8

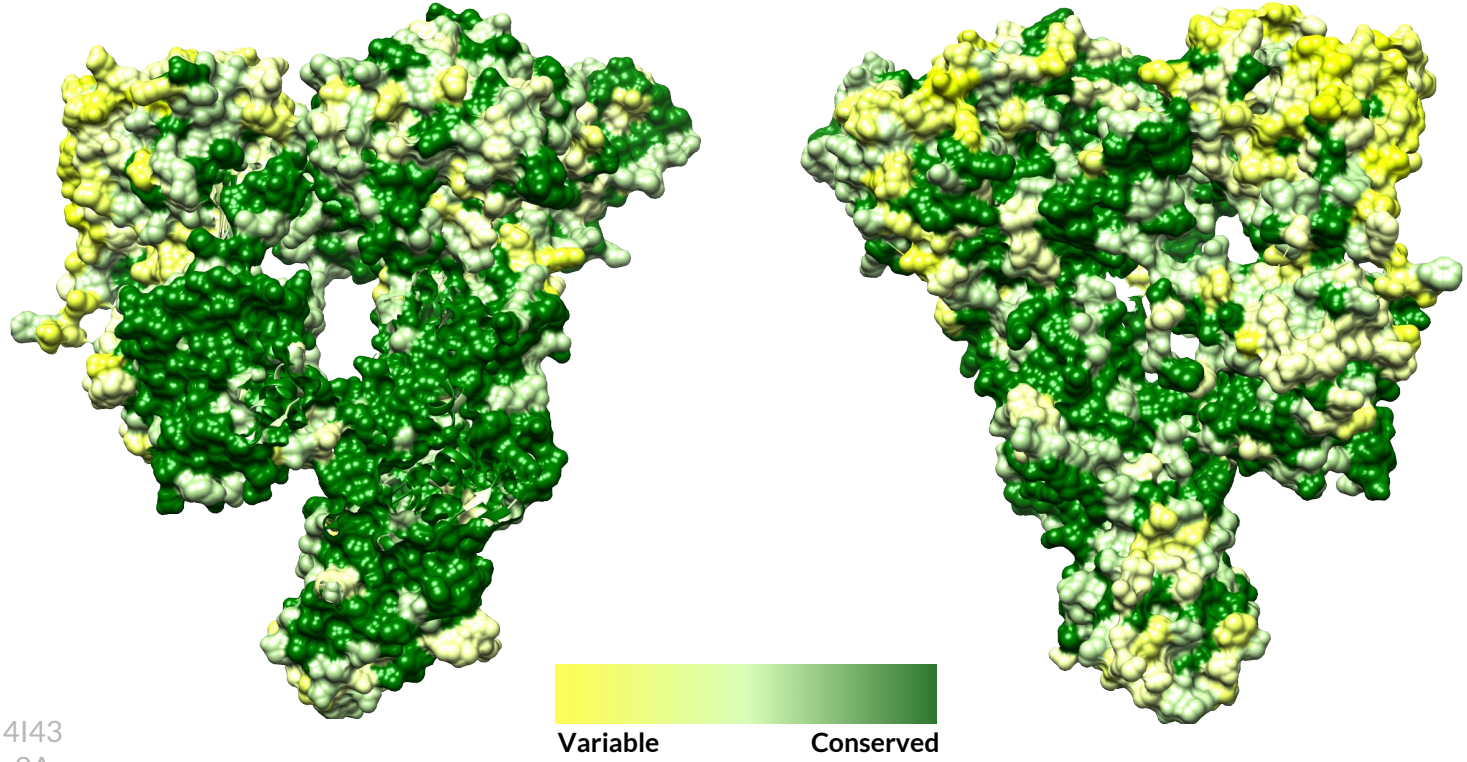
HMMalign from PFAM (of each domain)

```
                1261                                                                1320
H.sapiens       mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
M.mulatta       mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
C.lupus         mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
M.musculus      mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
G.gallus        mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
D.rerio         mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL
D.melanogaster  msgfecrilp kcrtqneeft HRDGVWNLQN EITKERTAQC FLRVDDESLG RFHNRVRQIL
C.elegans       msgfecrilp kcrtaneefv HRDGVWNLQN EVTKERTAQC FLKVDEESLS KFHNRIRQIL
S.cerevisiae    smcgfevril prqrmeevvs NDEGVWDLVD ERTKQRTAKA YLKVSEEEIK KFDSRIRGIL
S.pombe         mtgfevrilp kirqneefsl .KDGVWNLTD NRTKQRTAQA FIRVTEDGIN QFGNRIRQIL
N.crassa        cgfevrilpk irnqndefpv .KDSVWSLVD NTTKERTAHA FLQVTEEDIQ KFNNRIRQIL
A.thaliana      cgfevrvlpk irmgqeafss TRDGVWNLQN EQTKERTAVA FLRADDEHMK VFENRVRQIL
X.tropicalis    mcgfecrilp kcrtsyeeft HKDGVWNLQN EVTKERTAQC FLRVDDESMQ RFHNRVRQIL

                1321                                                                1380
H.sapiens       MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
M.mulatta       MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
C.lupus         MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
M.musculus      MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
G.gallus        MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
D.rerio         MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
D.melanogaster  MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
C.elegans       MSSGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
S.cerevisiae    MASGSTTFTK VAAKWNTSLI SLFTYFREAI VATEPLLDIL VKGETRIQNR VKLGLNSKMP
S.pombe         MSSGSTTFTK IANKWNTALI ALMTYYREAA ISTPELLDIL VKCESKIQTR VKISLNSKMP
N.crassa        MSSGSTTFTK IANKWNTALI ALFTYYREAA VSTVNLLDTI VKCETKIQTR VKIGLNSKMP
A.thaliana      MSSGSTTFTK IVNKWNTALI GLMTYFREAT VHTQELLDLL VKCENKIQTR VKIGLNSKMP
X.tropicalis    MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP

                1381                                                                1440
H.sapiens       SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
M.mulatta       SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
C.lupus         SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
M.musculus      SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
G.gallus        SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
D.rerio         SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
D.melanogaster  SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms hdedqlipnl
C.elegans       SRFPPVVFYT PKEIGGLGML SMGHVLIPQS DLRW.mqqte aggvthfrsg mshdedqlip
S.cerevisiae    TRFPPAVFYT PKELGGLGMI SASHILIPAS DLSWSkqtdt githfragmt hedeklipti
S.pombe         SRFPPAVFYS PKELGGLGML SMGHVLIPQS DLRWSkqtdt githfrsgmt tngehlipnl
N.crassa        SRFPPAVFYT PKELGGLGMI SGSHILIPTS DKRWSkqtdl gvthyragms hdeetlipni
A.thaliana      SRFPPVIFYT PKEIGGLGML SMGHILIPQS DLRYSnqtdv gvshfrsgms heedqlipnl
X.tropicalis    SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
```

ConSurf Color Coded MSA

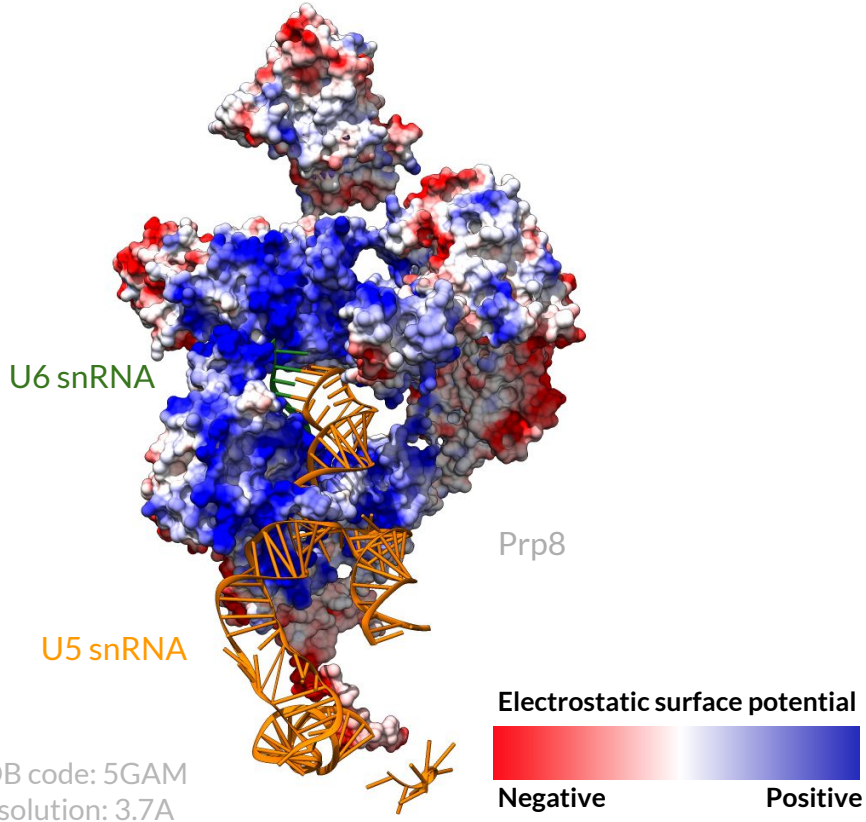


Prp8

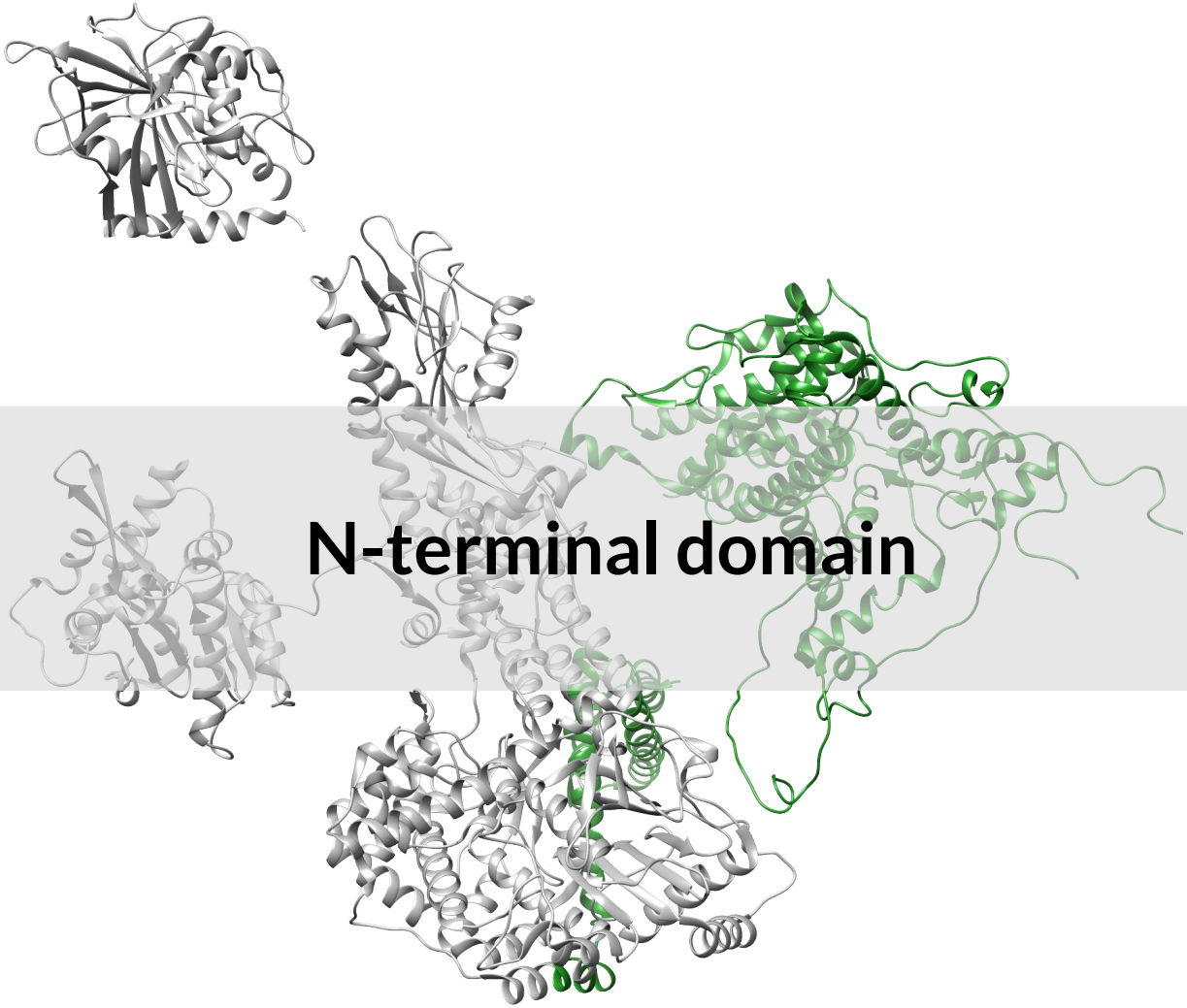


PDB code: 4I43
Resolution: 2A

Electrostatic surface potential in the cavity

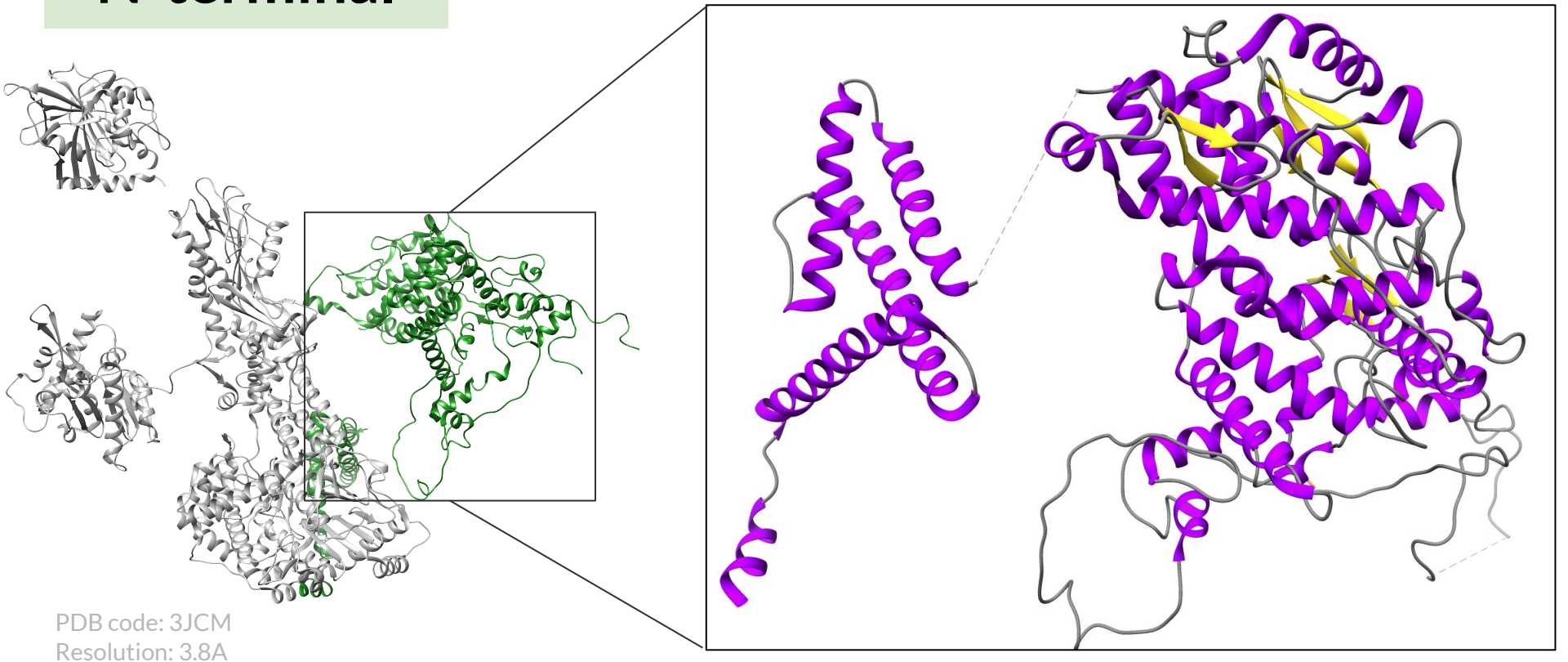


PDB code: 5GAM
Resolution: 3.7A



N-terminal domain

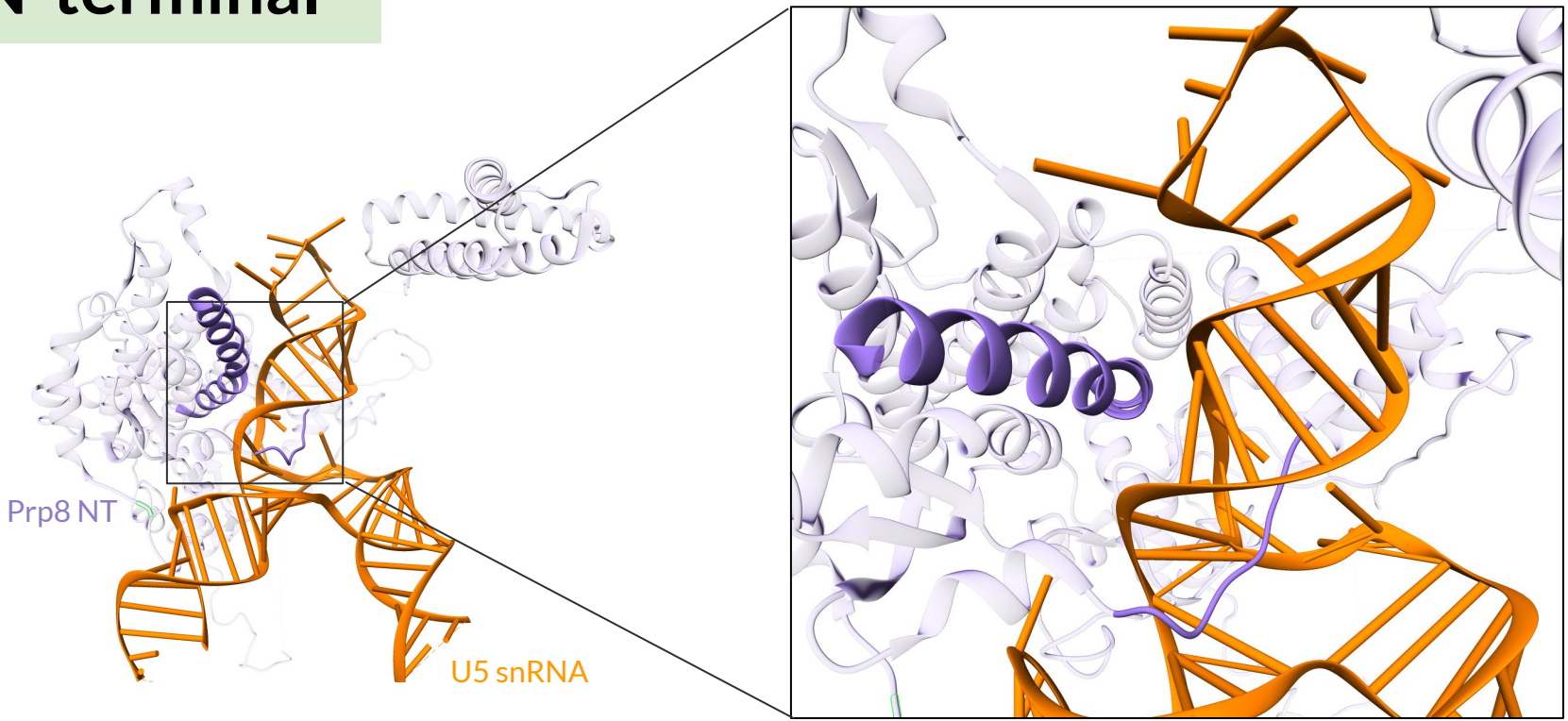
N-terminal



PDB code: 3JCM
Resolution: 3.8A

N-terminal

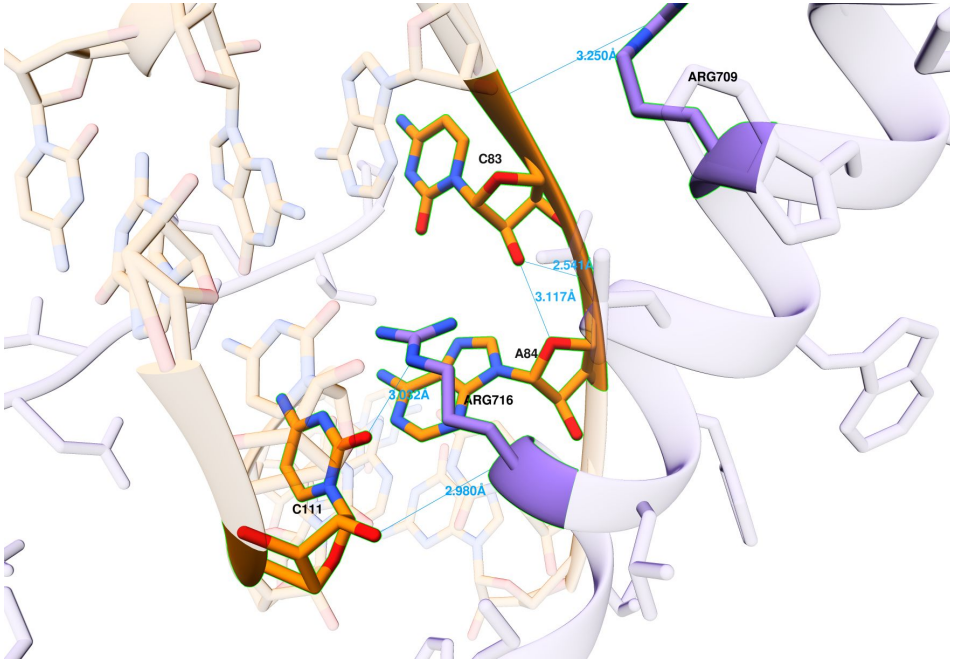
The N-domain recognises the U5 snRNA...



...and harnesses it with a polypeptide loop!

PDB code: 3JCM
Resolution: 3.8A

N-terminal



PDB code: 3JCM
Resolution: 3.8A

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

N-terminal

Negatively charged

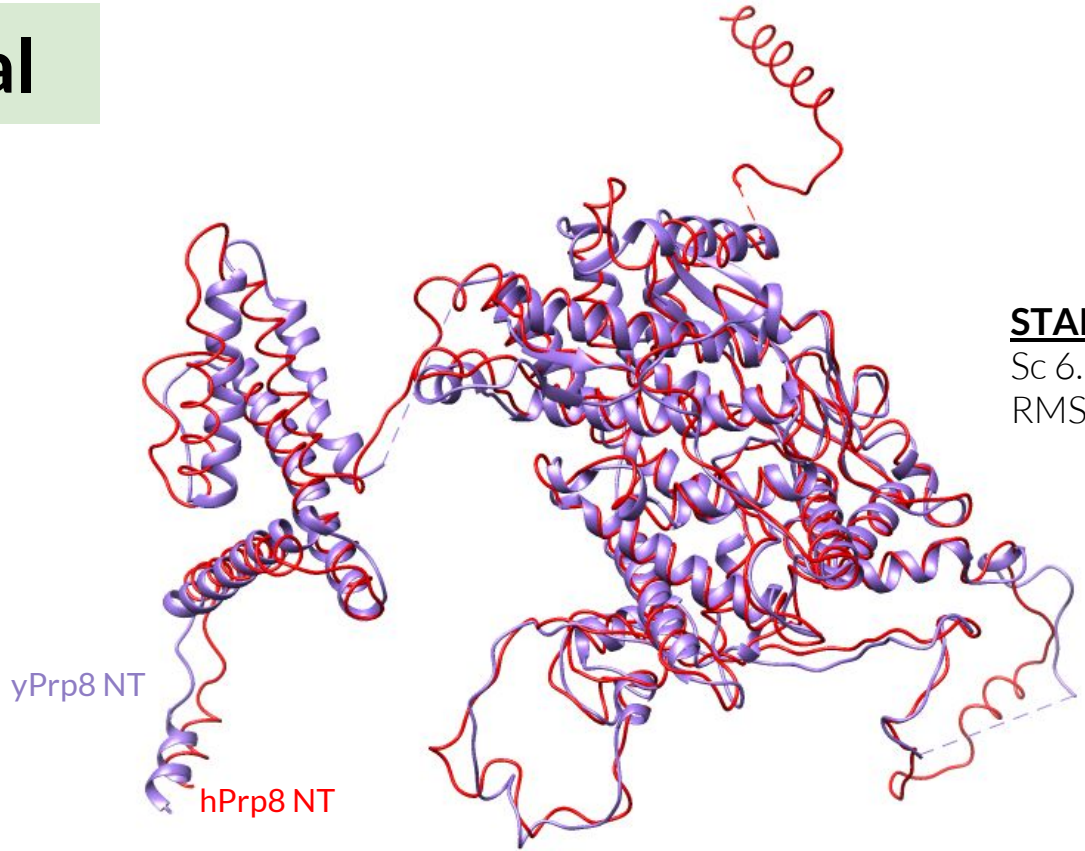
Positively charged

Non-polar

Polar

Sequence alignment based on PFAM domains

N-terminal

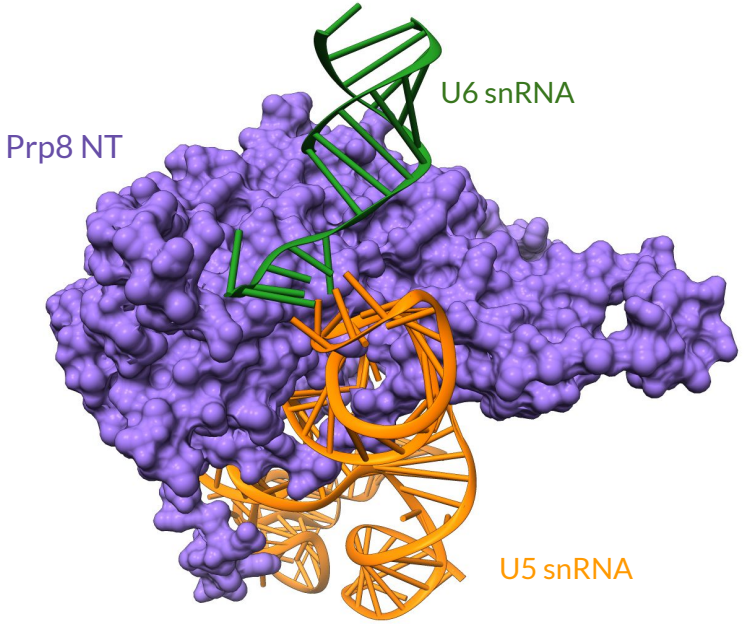


STAMP VALUES
Sc 6.53
RMS 1.58

PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A

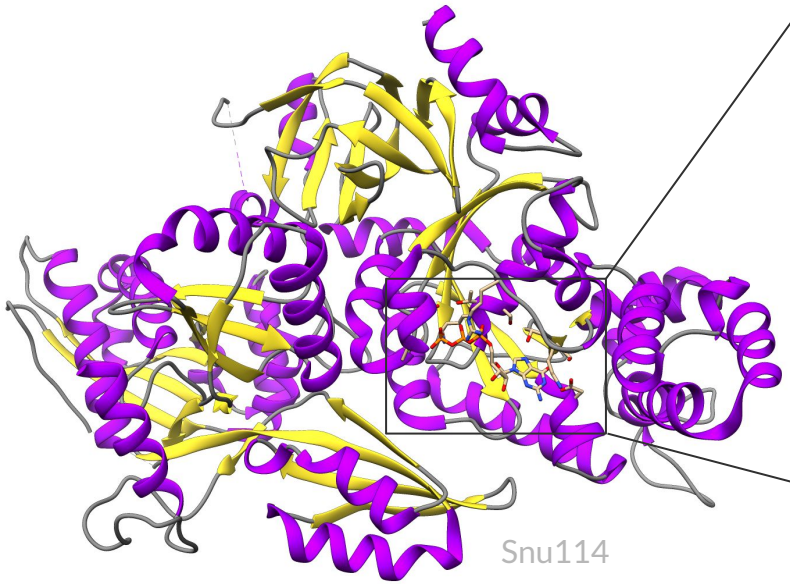
N-terminal



PDB code: 5GAM
Resolution: 3.7A

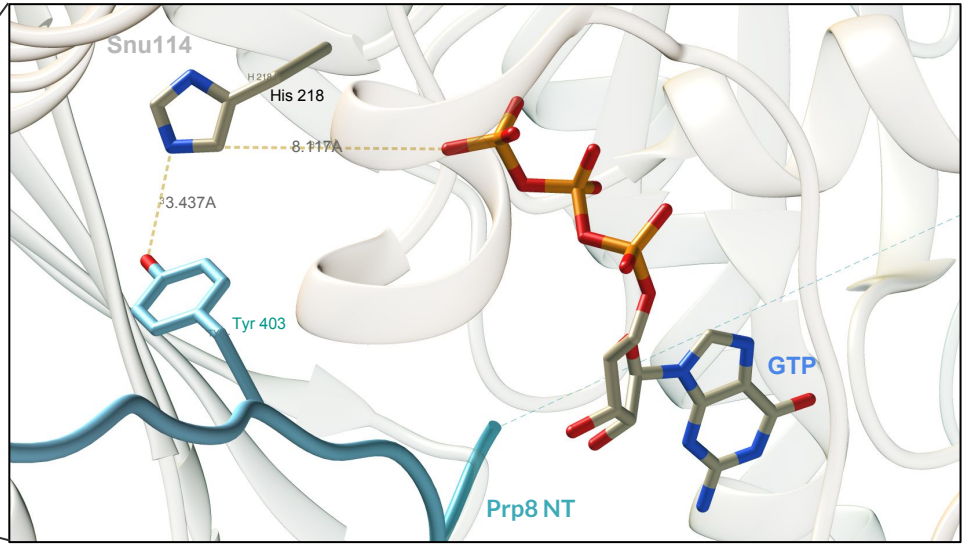
Snu114

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



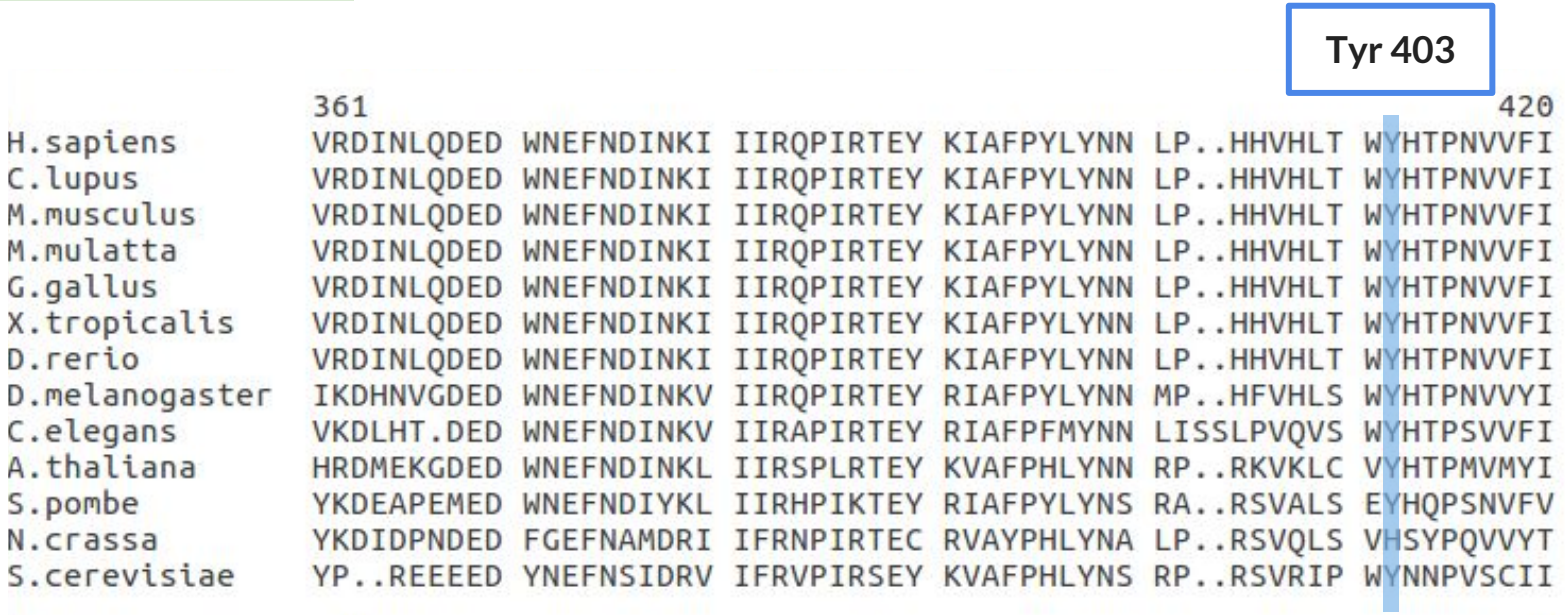
Snu114

PDB code: 5GAM
Resolution: 3.7A



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

N-terminal



Negatively charged

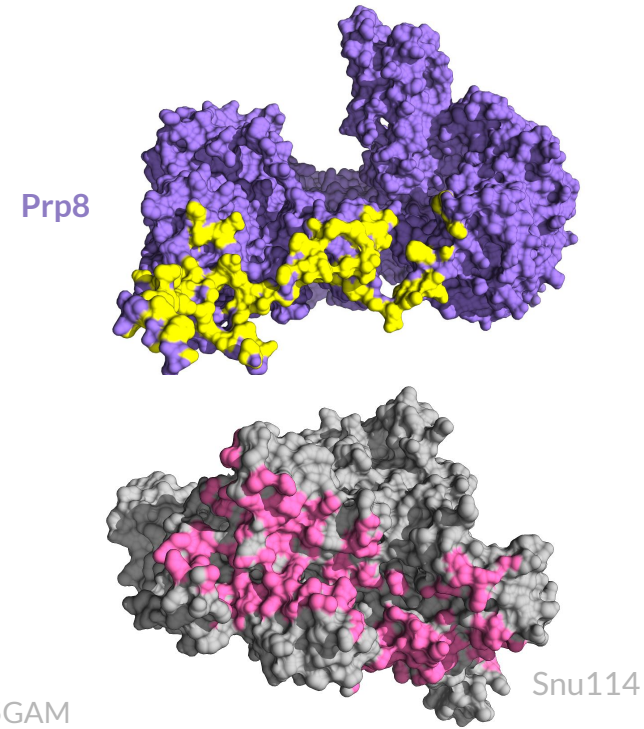
Non-polar

Positively charged

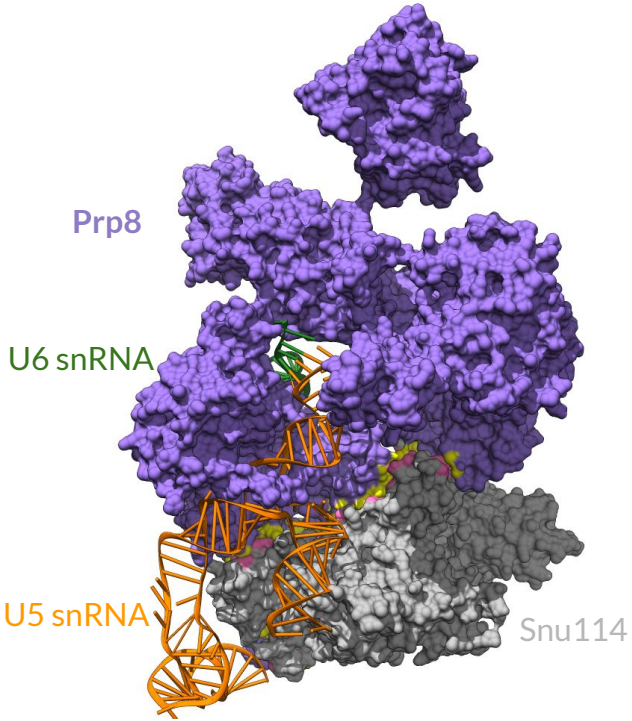
Polar

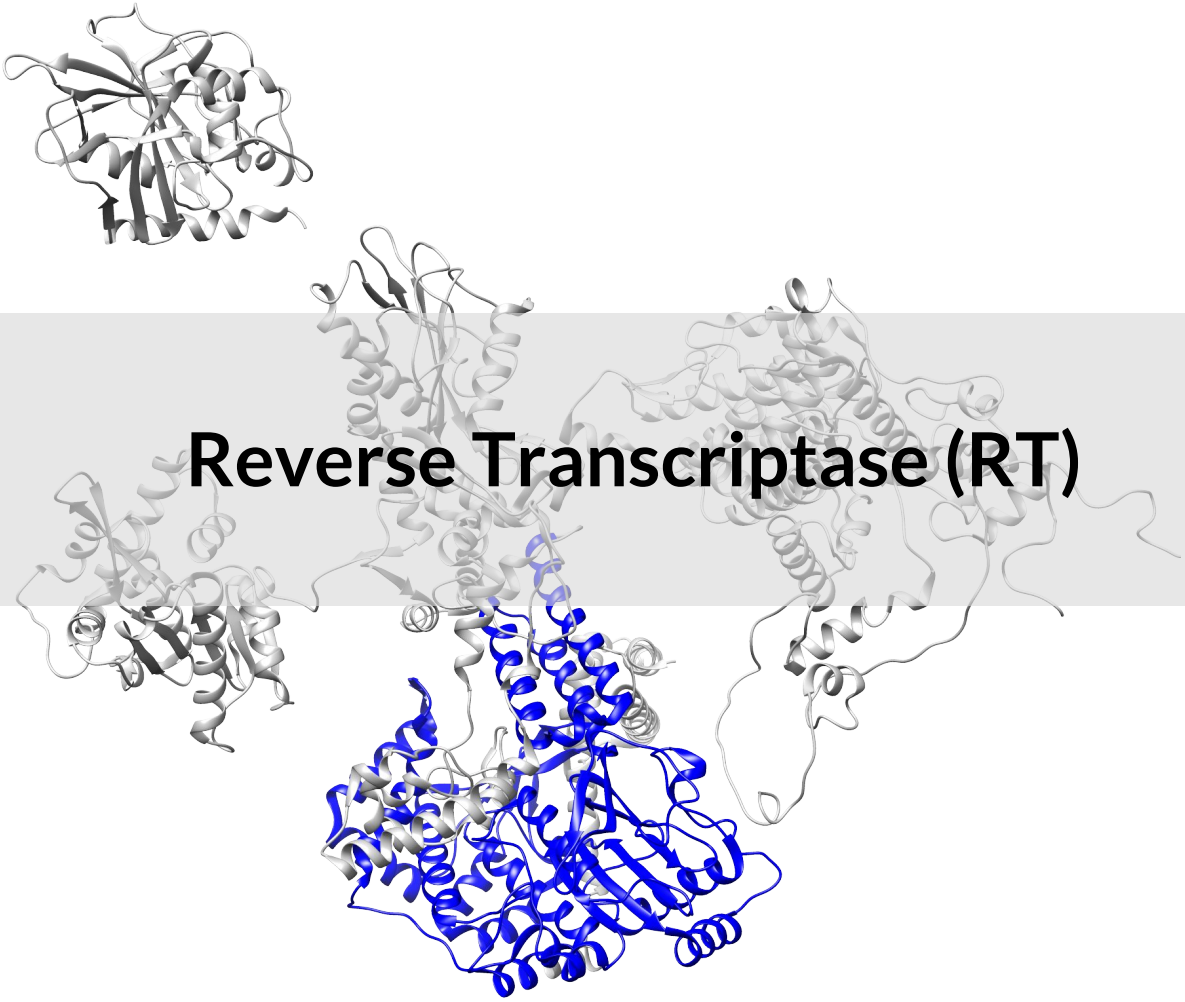
Clustal Sequence alignment

N-terminal - Snu114



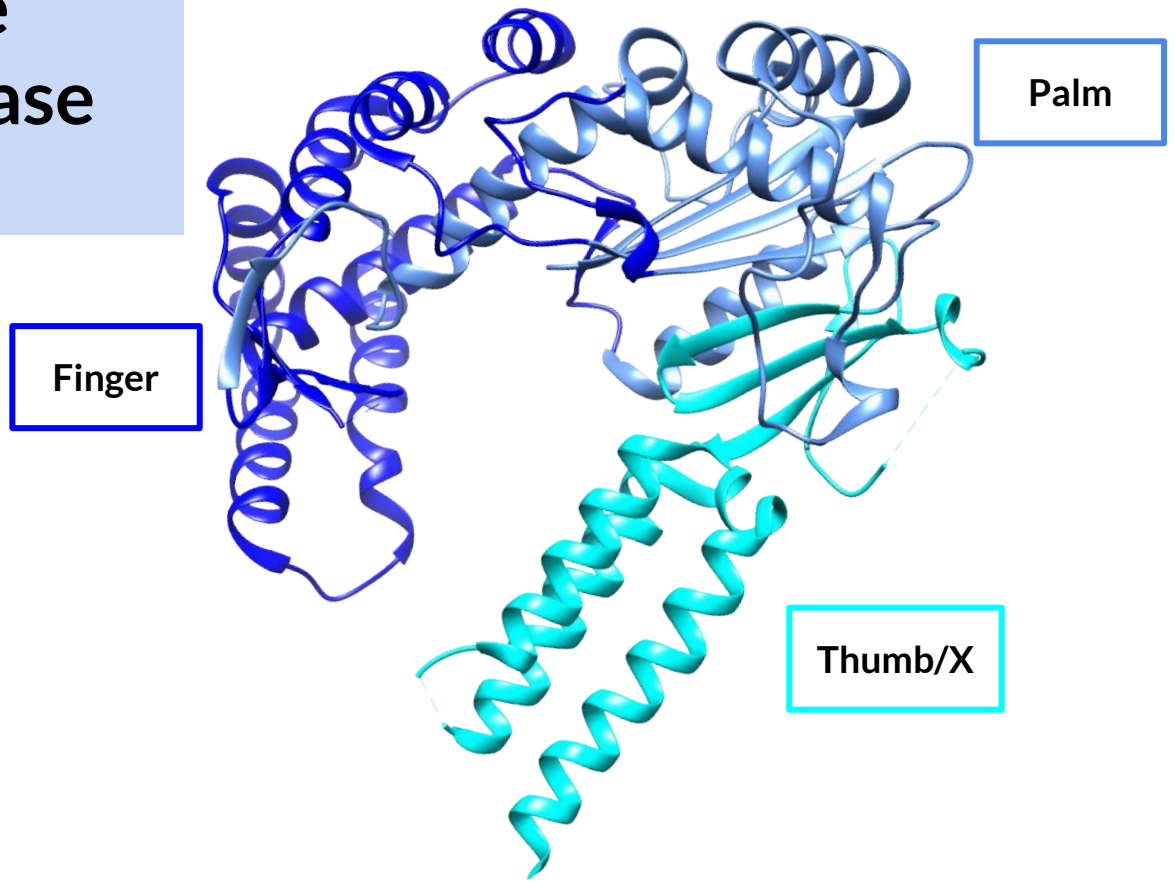
PDB code: 5GAM
Resolution: 3.7A





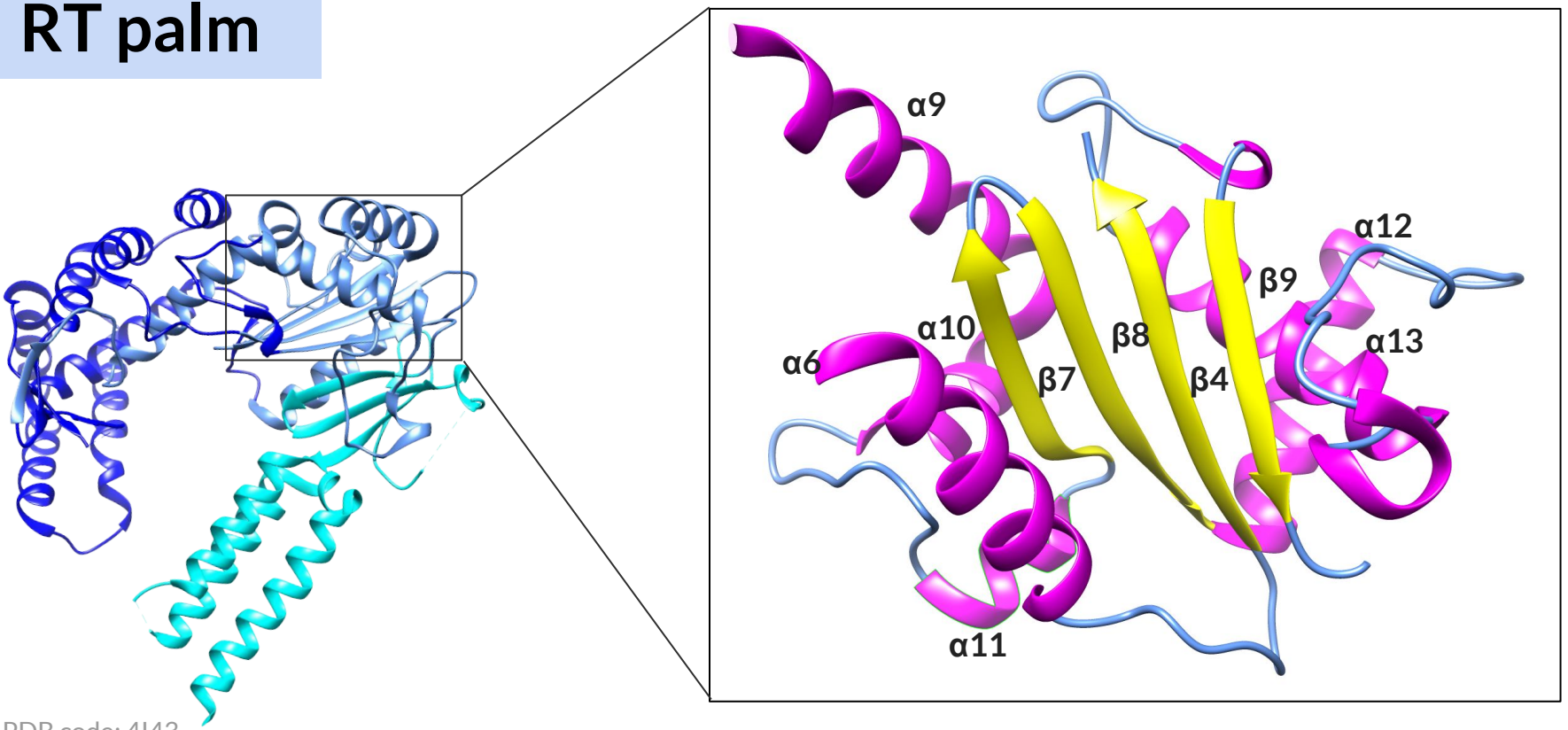
Reverse Transcriptase (RT)

Reverse transcriptase (RT)



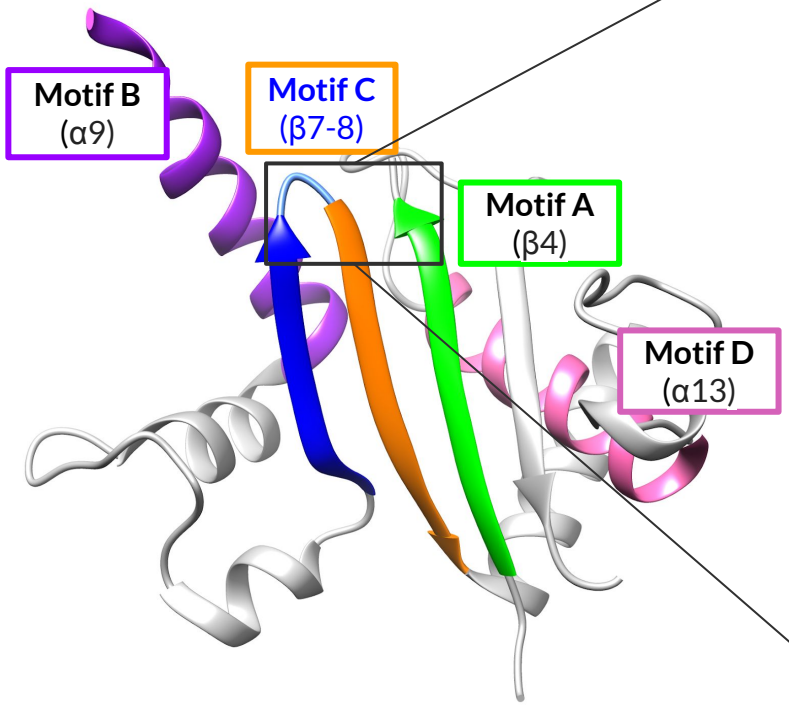
PDB code: 4I43
Resolution: 2A

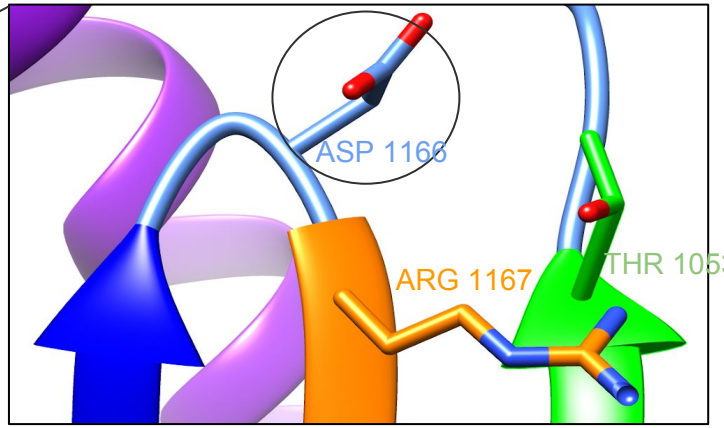
RT palm



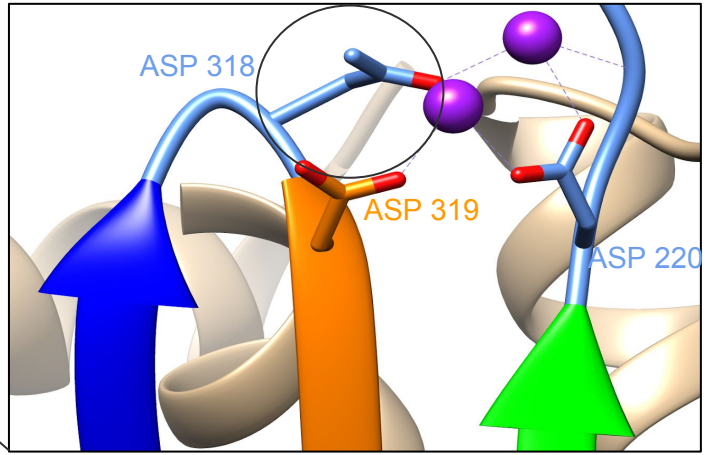
PDB code: 4I43
Resolution: 2A

RT palm





Prp8
PDB code: 4I43



HCV
PDB code: 1NB6

RT palm

Thr 1053

Motif A

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

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

RT palm

Asp 1166

Arg 1167

Motif C

	1141				1200
H.sapiens	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DIATEAAHPI	RLFCRYIDRI HIFFRFTADE
C.lupus	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DIATEAAHPI	RLFCRYIDRI HIFFRFTADE
M.musculus	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DIATEAAHPI	RLFCRYIDRI HIFFRFTADE
M.mulatta	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DIATEAAHPI	RLFCRYIDRI HIFFRFTADE
G.gallus	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DIATEVAHPI	RLFCRYIDRI HIFFRFTADE
X.tropicalis	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DVATESAHPI	RLFCRYIDRI HIFFRFSADE
D.rerio	LVMDLLVLGL	HRASEMAGPP	QMPNDFLSFQ	DTATESAHPI	RLYCRYIDRI HIFFRFSADE
D.melanogaster	LVLDLLVLGL	HRSSEMAGPP	QMPNDFLTFQ	DTVTETAHPI	RLYCRYVDRI HLFFRFSAEE
C.elegans	LVLDLLVLGL	RRASEIAGPP	QCPNEFLQFQ	DVATEIGHPI	RLYCRYIDRV WIMFRFSADE
A.thaliana	LVLDLLLLGL	TRASEIAGPP	QRPNEFMTYW	DTKVETRHPI	RLYSRYIDKV HIMFKFTHEE
S.pombe	LVLDLLILGL	QRATEIAGPA	DAPNDFLHFK	DQATETSHPI	RLYTRYIDKV YIMFRFTDEE
N.crassa	LILDLLLLGP	QRASEIAGPP	HAPNDFLQFK	DRETETRHPI	RLYTRYIDKI WVFLRFTADE
S.cerevisiae	LVIDLLLLGQ	ERATDLAGPA	NNPNEFMQFK	SKEVEKAHPI	RLYTRYLDRI YMLFHFEEDE

Negatively charged

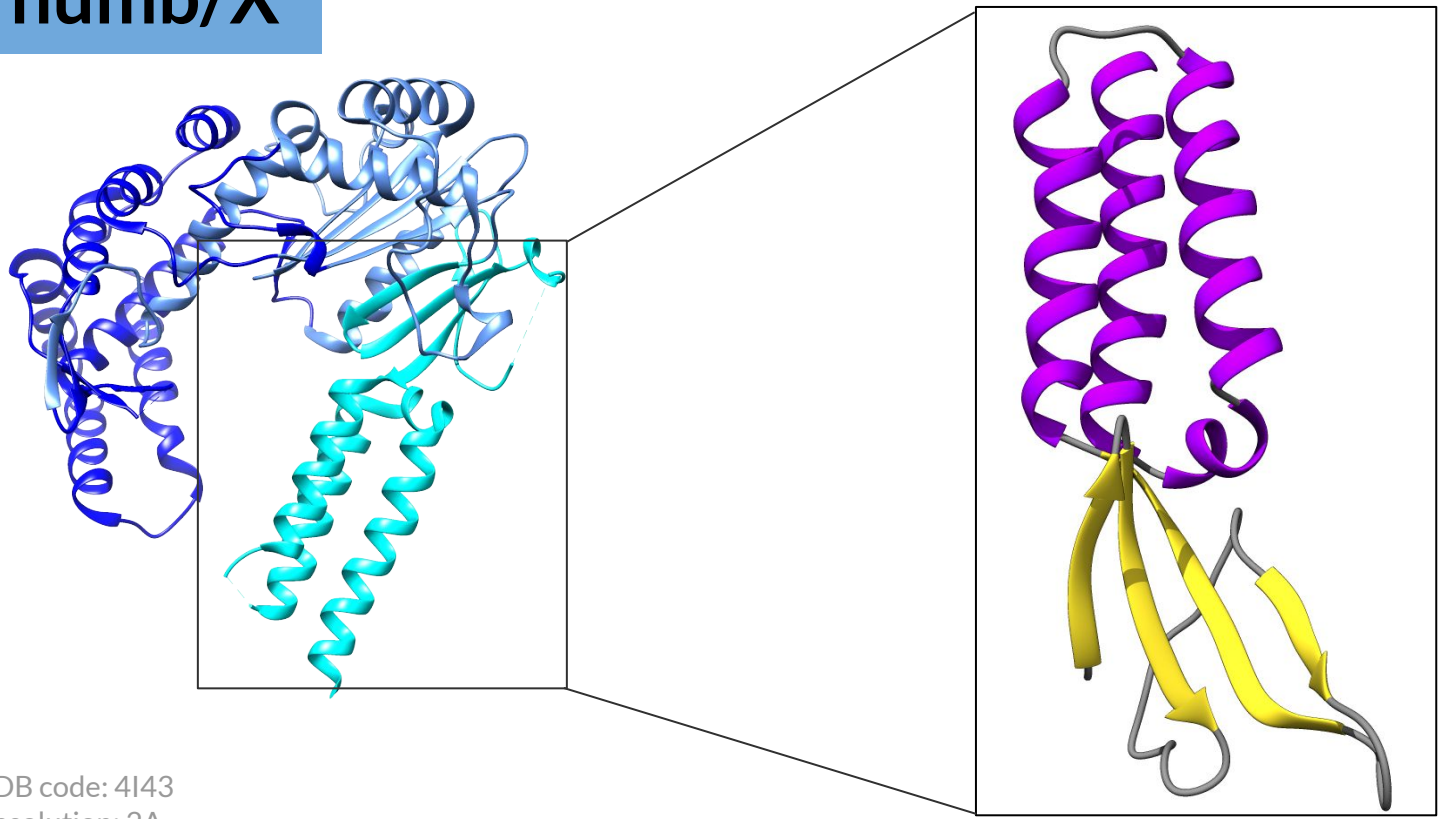
Non-polar

Positively charged

Polar

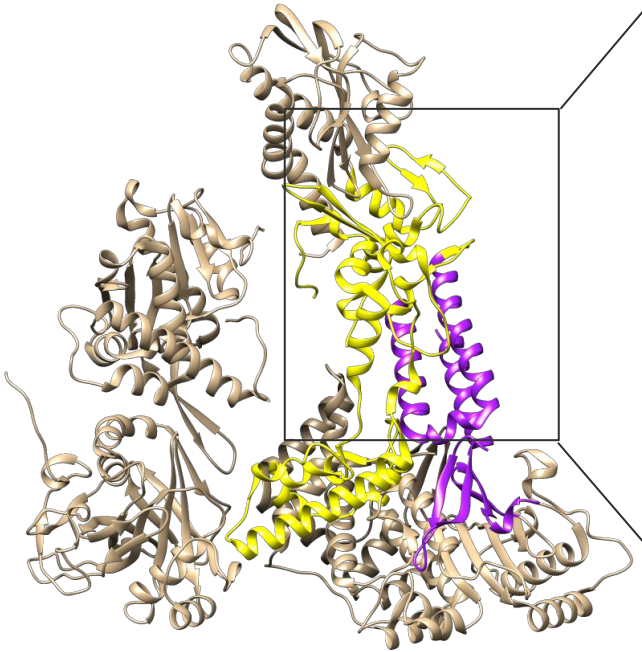
Clustal Sequence alignment

Thumb/X

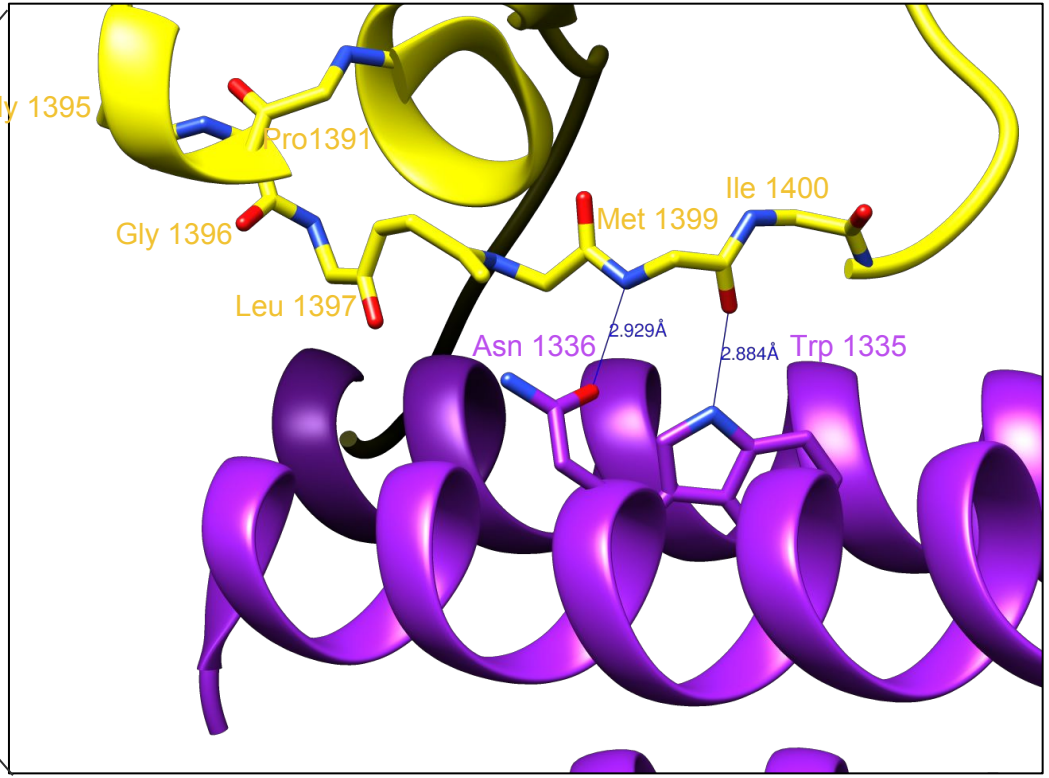


PDB code: 4I43
Resolution: 2A

Thumb/X - Linker



PDB code: 4I43
Resolution: 2A



Linker Thumb/X

Thumb/X - Linker

Trp 1335

Asn 1336

Thumb region

```
                1321                                                                1380
H.sapiens       MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
M.mulatta       MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
C.lupus         MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
M.musculus      MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
G.gallus        MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
D.rerio         MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
D.melanogaster  MASGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
C.elegans       MSSGSTTFTK IVNKWNTALI GLMTYFREAV VNTQELLDLL VKCENKIQTR IKIGLNSKMP
S.cerevisiae    MASGSTTFTK VAAKWNTSLI SLFTYFREAI VATEPLLDIL VKGETRIQNR VKLGLNSKMP
S.pombe         MSSGSTTFTK IANKWNTALI ALMTYYREAA ISTPELLDLL VKCESKIQTR VKISLNSKMP
N.crassa        MSSGSTTFTK IANKWNTALI ALFTYYREAA VSTVNLLDTI VKCETKIQTR VKIGLNSKMP
A.thaliana      MSSGSTTFTK IVNKWNTALI GLMTYFREAT VHTQELLDLL VKCENKIQTR VKIGLNSKMP
X.tropicalis    MASGSTTFTK IVNKWNTALI 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                                                    1440
H.sapiens       SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
M.mulatta       SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
C.lupus         SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
M.musculus      SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
G.gallus        SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
D.rerio         SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
D.melanogaster  SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms hdedqlipnl
C.elegans       SRFPPVVFYT PKEIGGLGML SMGHVLIPQS DLRW.mqqte aggvthfrsg mshdedqlip
S.cerevisiae    TRFPPAVFYT PKELGGLGMI SASHILIPAS DLSWSkqtdt githfragmt hedeklipti
S.pombe         SRFPPAVFYS PKELGGLGML SMGHVLIPQS DLRWSkqtdt githfrsgmt tngehlipnl
N.crassa        SRFPPAVFYT PKELGGLGMI SGSHILIPTS DKRWSkqtdl gvthyragms hdeetlipni
A.thaliana      SRFPPVIFYT PKEIGGLGML SMGHILIPQS DLRYSnqtdv gvshfrsgms heedqlipnl
X.tropicalis    SRFPPVVFYT PKELGGLGML SMGHVLIPQS DLRWSkqtdv githfrsgms heedqlipnl
```

Pro 1384
Pro 1385
Pro 1391

Gly 1398
Met 1399
Leu/Ile 1400

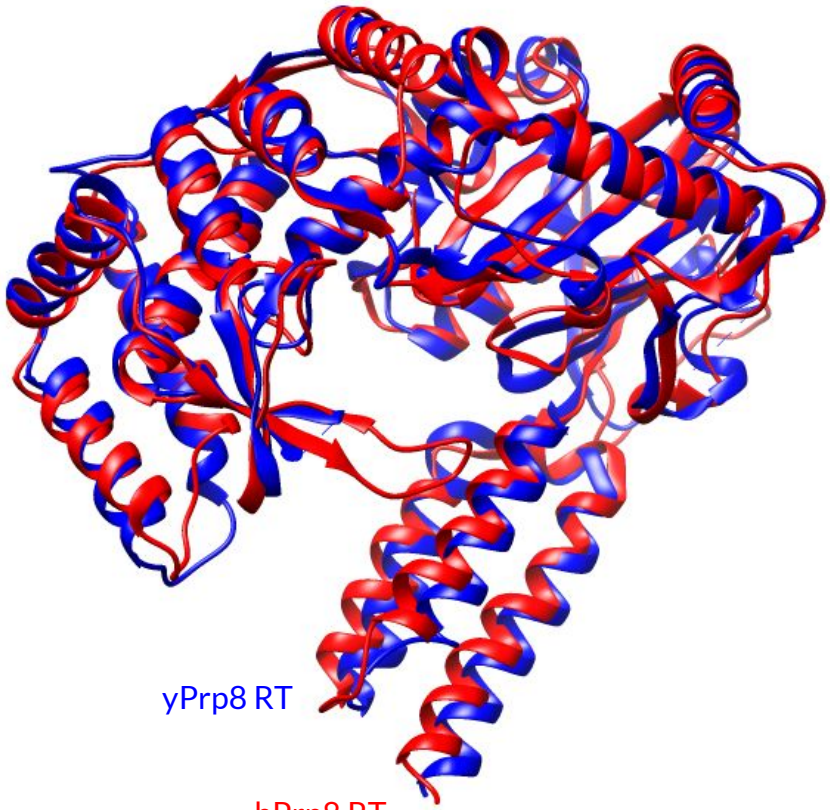
Negatively charged
Non-polar
Positively charged
Polar

Reverse transcriptase (RT)

STAMP VALUES

Sc 7.21

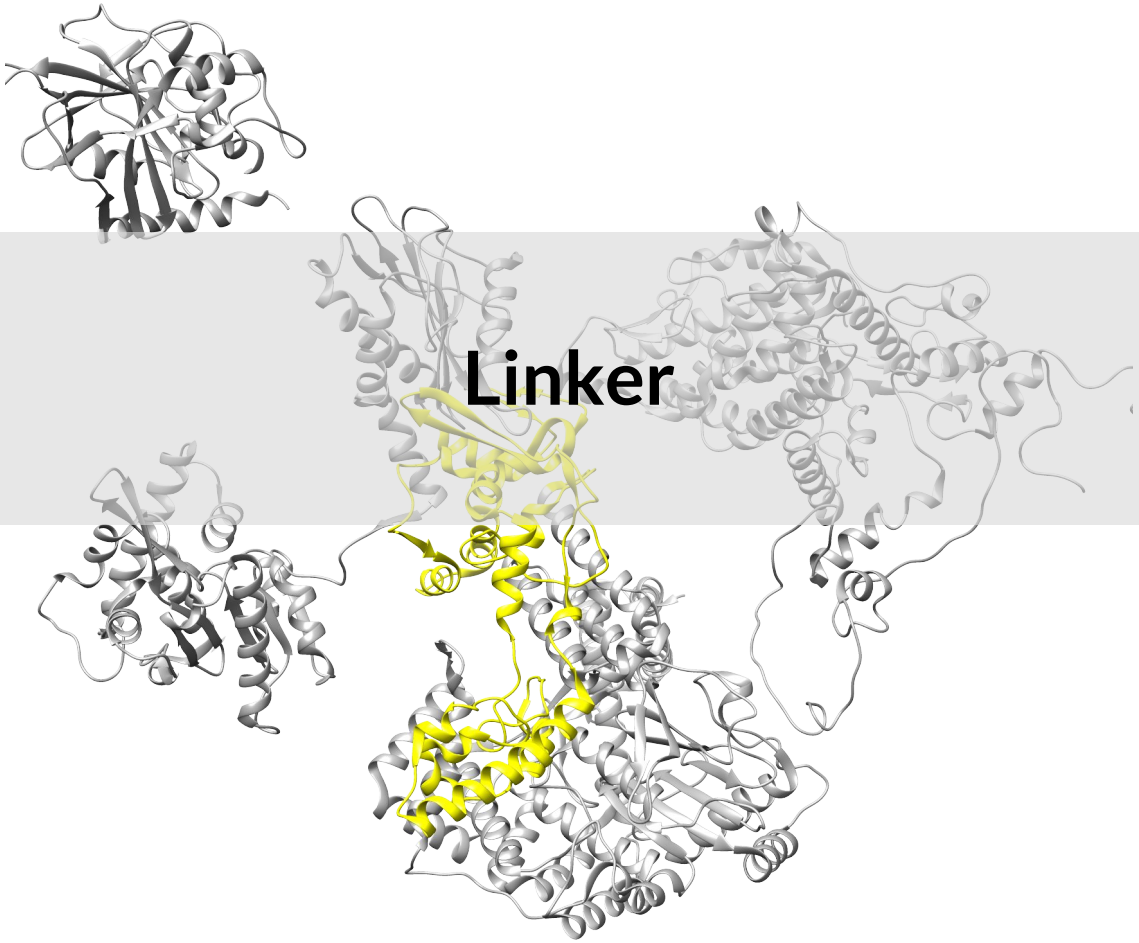
RMS 1.65



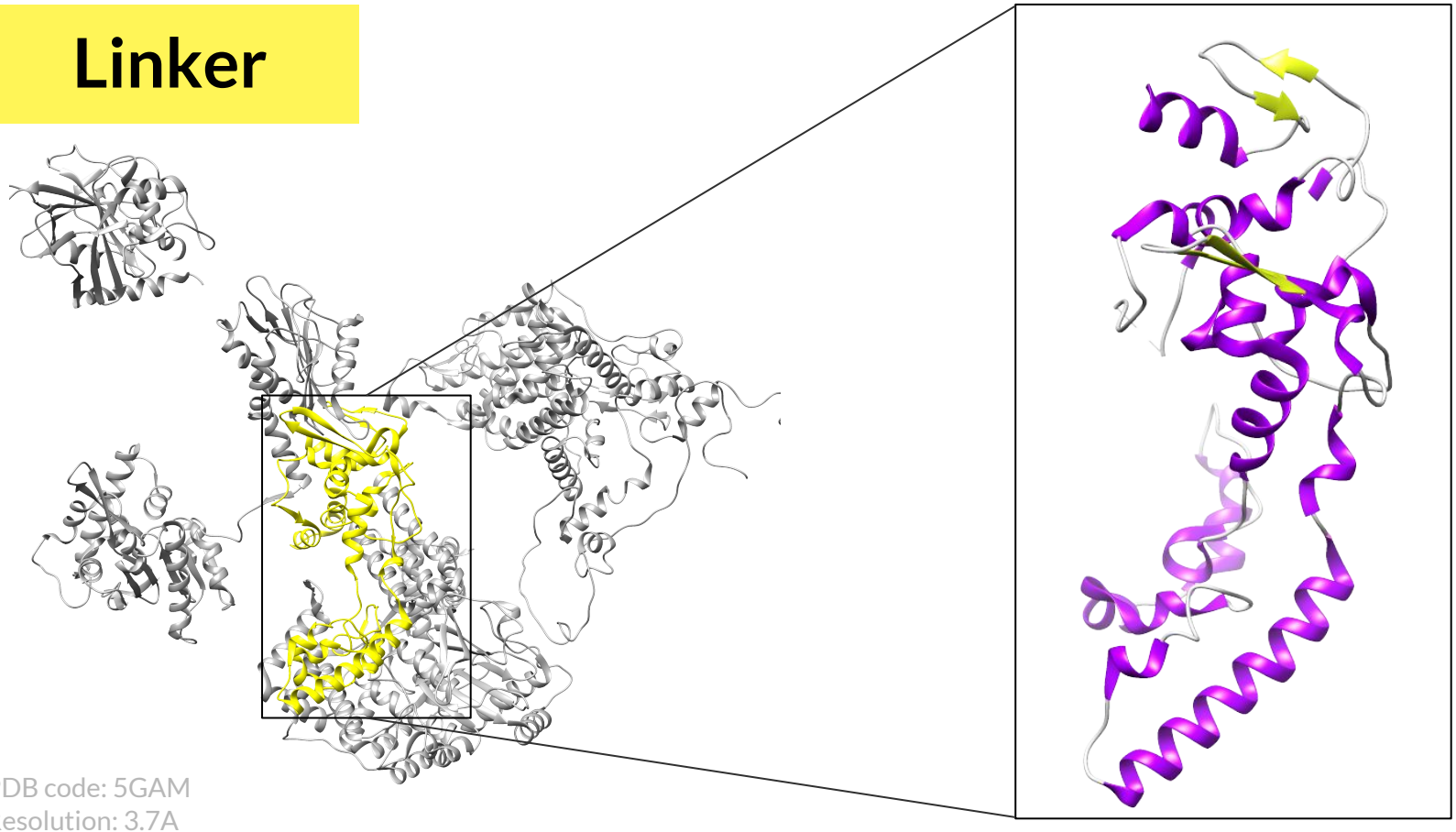
PDB code: 5MQF
Resolution: 5.9A

PDB code: 4I43
Resolution: 2A

Linker

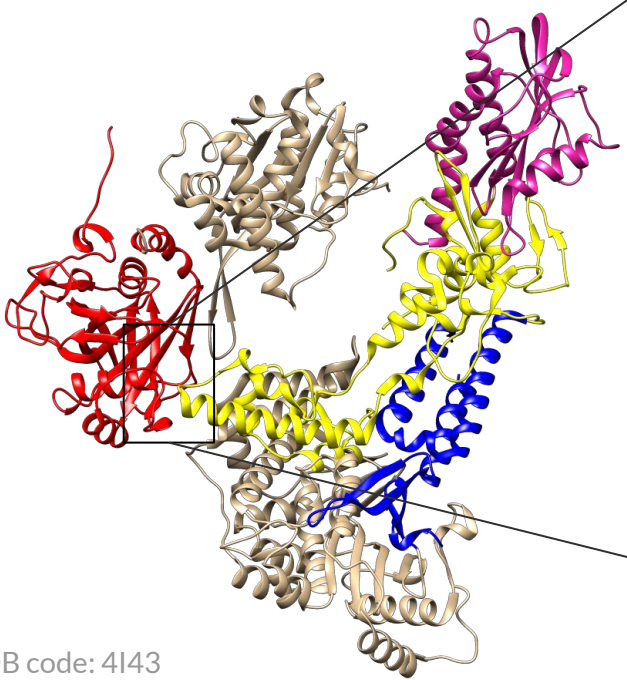


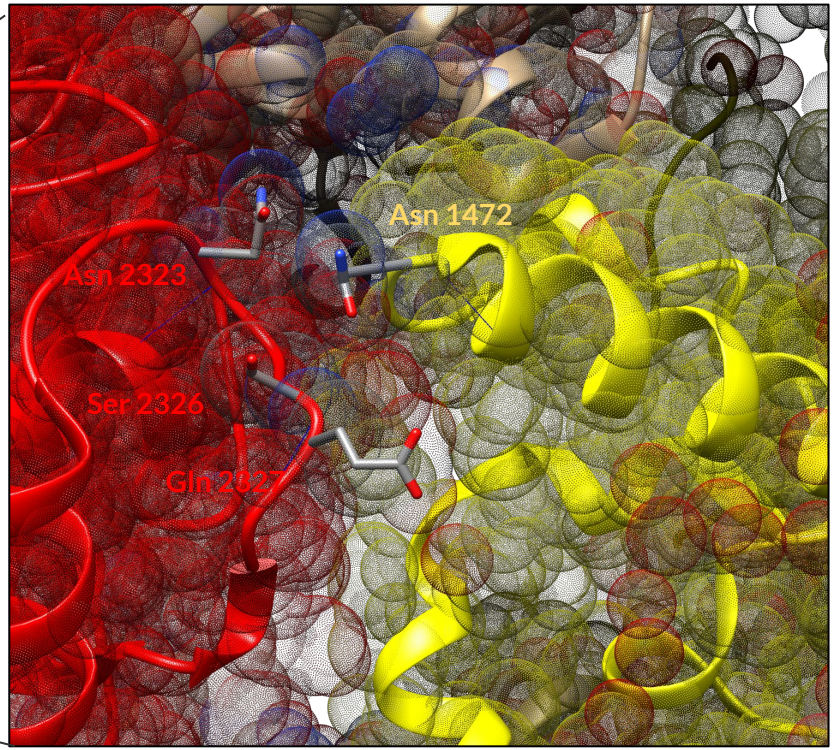
Linker



PDB code: 5GAM
Resolution: 3.7A

Linker- Jab1/MPN





Linker
Jab1/MPN
Thumb/X
Endonuclease

PDB code: 4I43
Resolution: 2A

Linker- Jab1/MPN

Jab1/MPN region

Asn 2323
Ser 2326

Glu 2327

	2281					2340
H.sapiens	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
M.mulatta	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
C.lupus	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
M.musculus	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
G.gallus	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
D.rerio	pswdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML
D.melanogaster	snwdgekTIV	ITCSFTPGSC	SLTAYKLTPS	GFEWGSKNTD	.KGNNPKGYL	PSHYERVQML
C.elegans	iswdgekTVM	ITCSFTPGSV	SLTAYKLTPS	GYEWGKANTD	.KGNNPKGYM	PTHYEKVQML
S.cerevisiae	fadkkrdCID	ISIFSTPGSV	SLSAYNLTDE	GYQWGEENKD	IMNVLSEGFE	PTFSTHAQLL
S.pombe	pewdtka.VT	LTVSYIPGSI	SLAAYTVSKE	GIEWGSKNMD	INSDEAIGYE	PSMAEKCQLL
N.crassa	nwdkqnt.LT	VAVSFTPGSV	SLSAWALTPQ	GFKWGVENKD	IASDQPQGFT	TSMGEKRQLL
A.thaliana	kqwdaekCII	LTCSFTPGSC	SLTSYKLTQA	GYEWGRLNKD	.TGSNPHGYL	PTHYEKVQML
X.tropicalis	pawdgekTII	ITCSFTPGSC	TLTAYKLTPS	GYEWGRQNTD	.KGNNPKGYL	PSHYERVQML

Negatively charged
Non-polar
Positively charged
Polar

Sequence alignment based on PFAM domains

Linker- Jab1/MPN

Linker region

Asn 1472

```
                1441                                                                1500
H.sapiens       SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
C.lupus         SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
M.musculus      SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
M.mulatta       SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
G.gallus        SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
X.tropicalis    SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
D.rerio         SKQTDVG.IT HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAIAQNR
D.melanogaster  SKQTDVG.IT HFRSGMSHDE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEANAQNR
C.elegans       MQQTEAGGVT HFRSGMSHDE DQLIPNLYRY IQPWEAEFVD SVRVWAEYAL KRQEANAQNR
A.thaliana      SNQTDVG.VS HFRSGMSHEE DQLIPNLYRY IQPWESEFID SQRVWAEYAL KRQEAQAQNR
S.pombe         SKQTDTG.IT HFRSGMTTNG EHLIPNLYRY IQPWESEFID SQRVWAEYAM KRQEALQQNR
N.crassa        SKQTDLG.VT HYRAGMSHDE ETLIPNIFRY IIPWEAEFID SQRVWTEYSQ KRLEANQQNR
S.cerevisiae    SKQTDTG.IT HFRAGMTHED EKLIPTIFRY ITTWENEFLD SQRVWAEYAT KRQEAIQQNR
```

Negatively charged

Non-polar

Positively charged

Polar

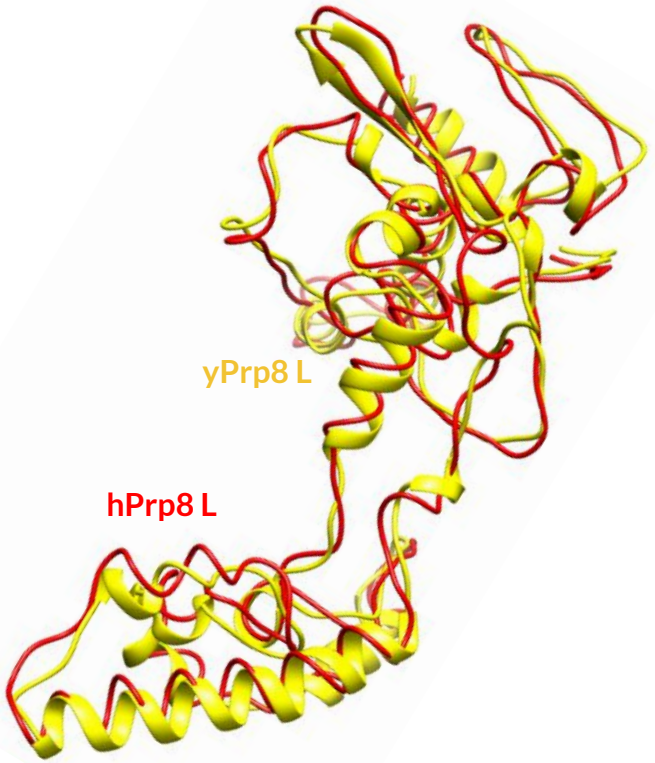
Clustal Sequence alignment

Linker

STAMP VALUES

Sc 8.43

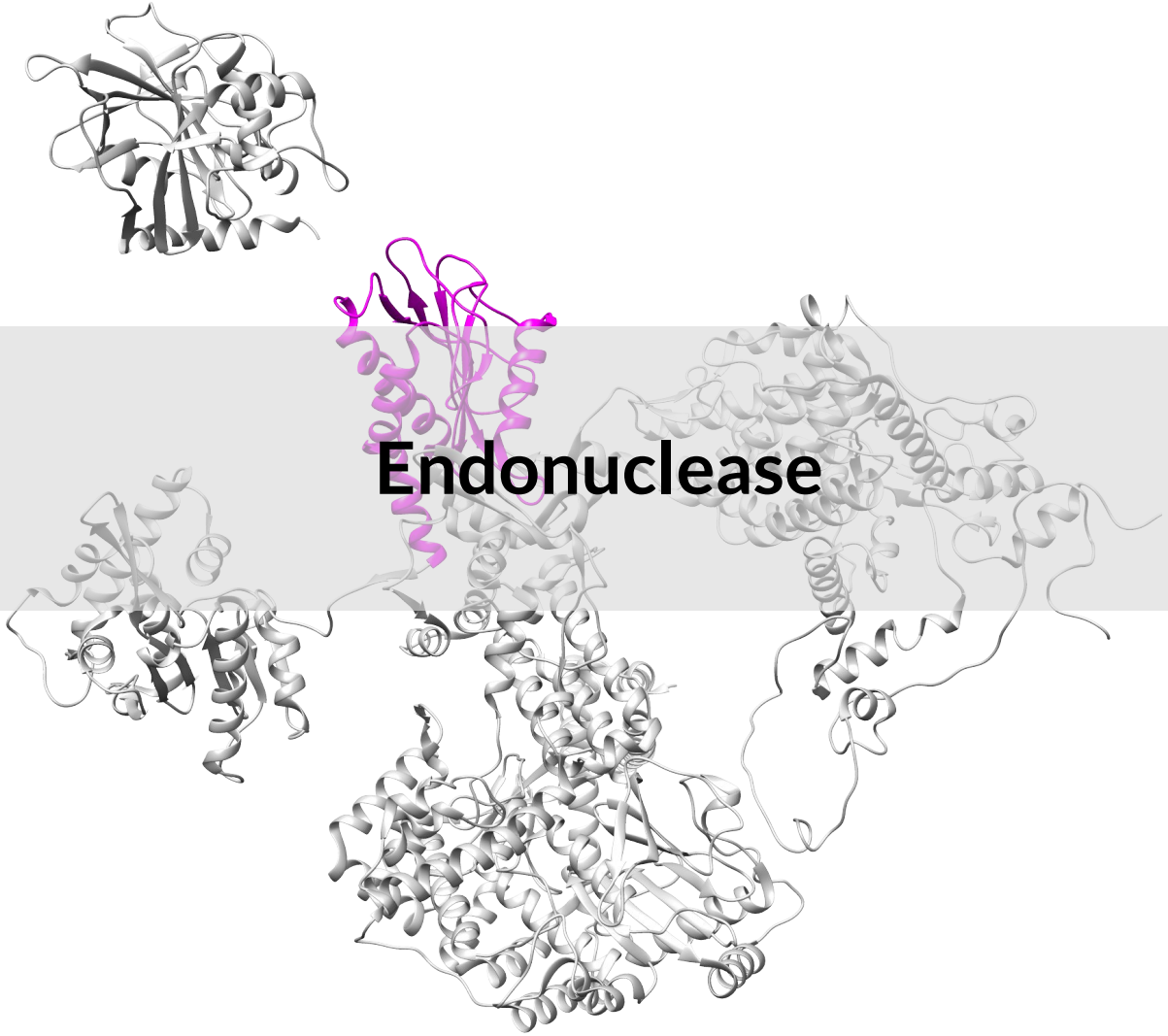
RMS 1.42



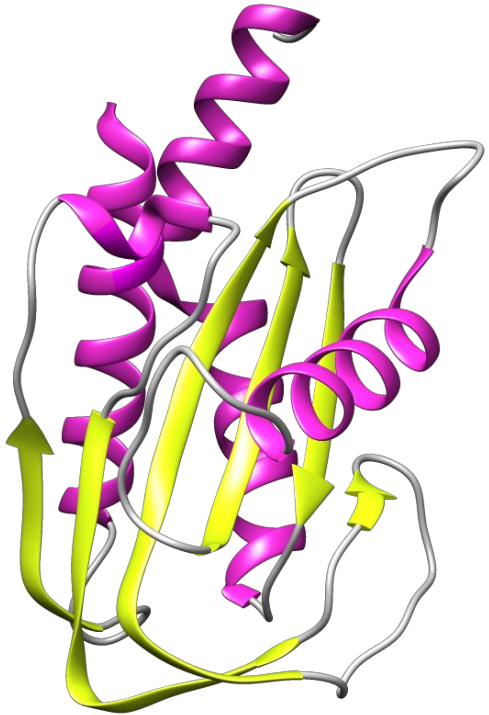
PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A

Endonuclease

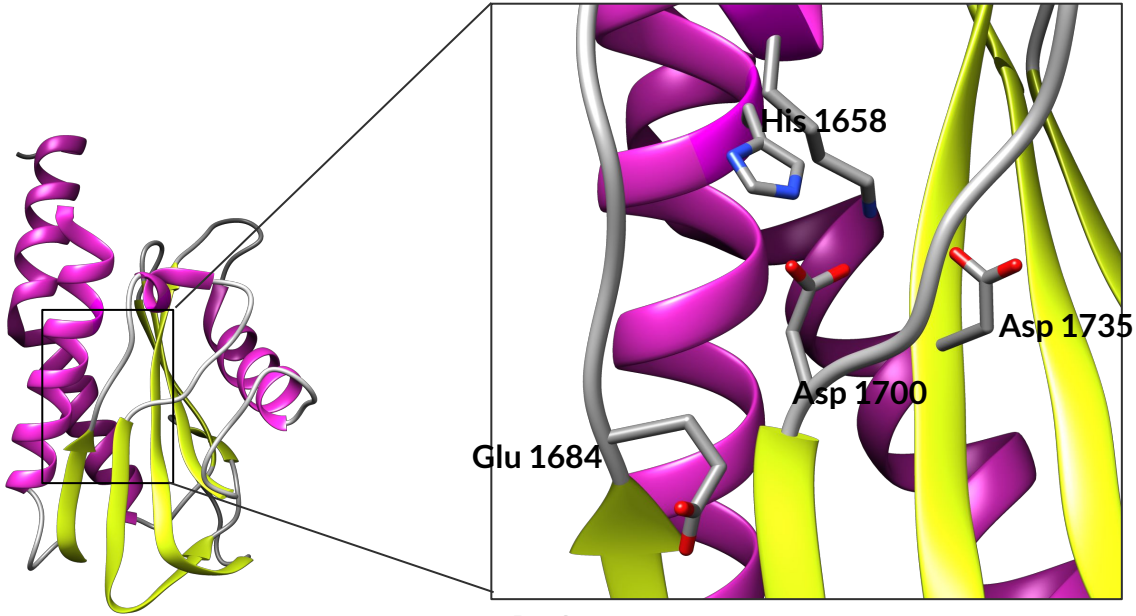


Endonuclease

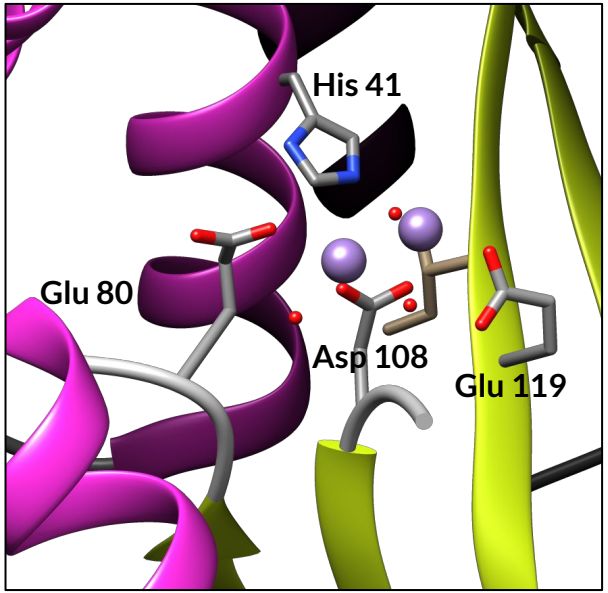


PDB code: 4I43
Resolution: 2A

Endonuclease

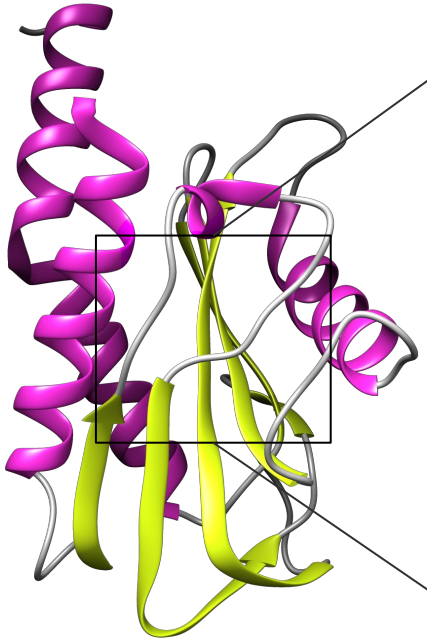


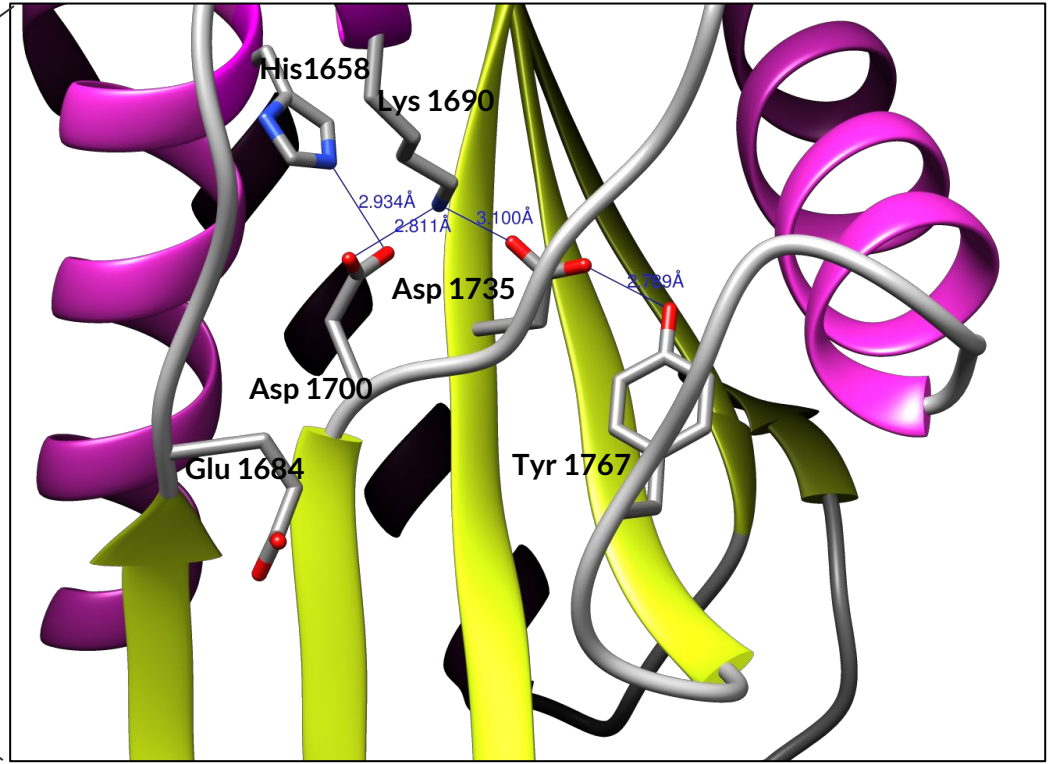
Prp8
PDB code: 4I43



Influenza Virus
PDB code: 2W69

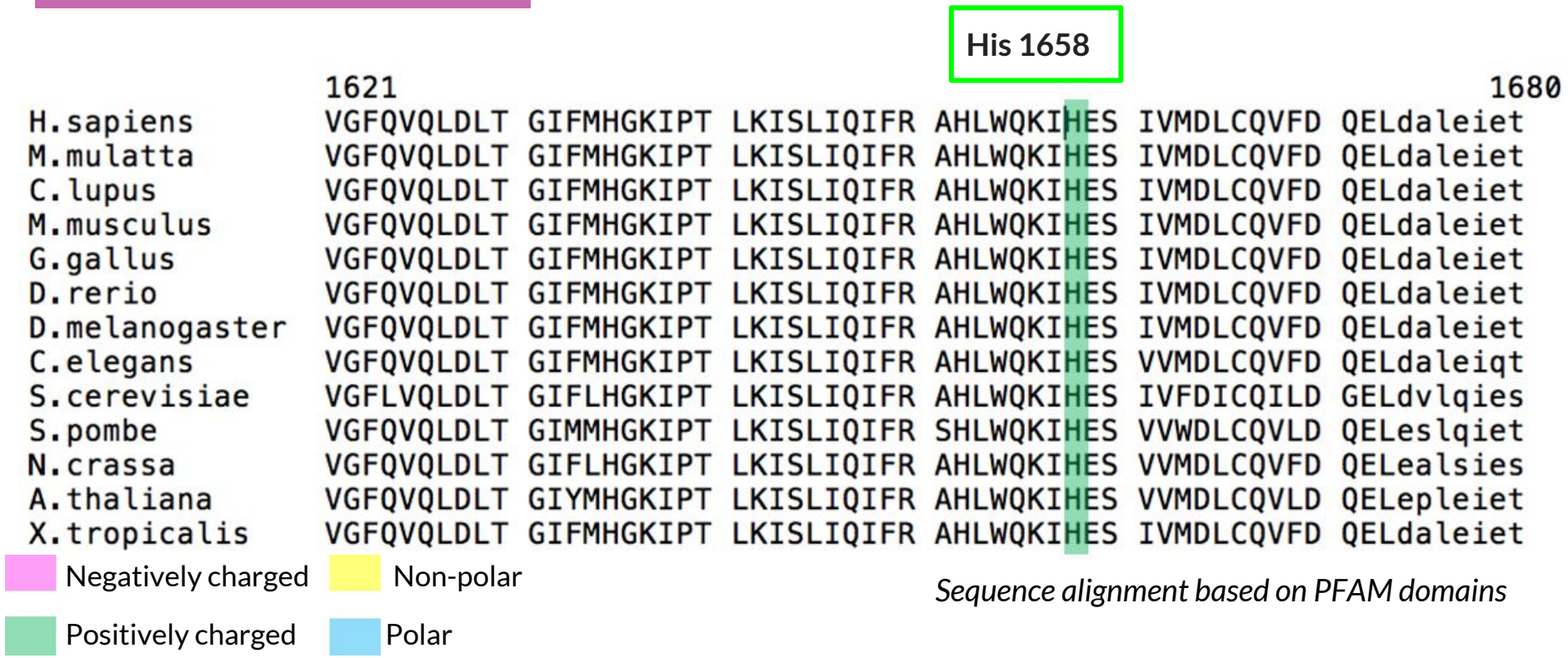
Endonuclease





PDB code: 4I43
Resolution: 2A

Endonuclease



Sequence alignment based on PFAM domains

Endonuclease

Glu 1684 | Lys 1690 | Asp 1700

```
                1681                                                                1740
H.sapiens       WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
C.lupus         WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
M.musculus      WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
M.mulatta       WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
G.gallus        WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
X.tropicalis    WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
D.rerio         WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
D.melanogaster  WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FPAYKWNVSR
C.elegans       WQKIHESVVM DLCQVFDQEL DALEIQTVQK ETIHPRKSYK MNSSCADVLL FAQYKWNVSR
A.thaliana      WQKIHESVVM DLCQVLDQEL EPLEIETVQK ETIHPRKSYK MNSSCADVLL FAAHKWPMSK
S.pombe         WQKIHESVVW DLCQVLDQEL ESLQIETVQK ETIHPRKSYK MNSSCADILL LAAYKWNVSR
N.crassa        WQKIHESVVM DLCQVFDQEL EALSIESVQK ETIHPRKSYK MNSSCADIQL FASHKWNVTR
S.cerevisiae    WQKIHESIVF DICQILDGEL DVLQIESVTK ETVHPRKSYK MNSSAADITM ESVHEWEVSK
```

Negatively charged | Non-polar

Positively charged | Polar

Clustal Sequence alignment

Endonuclease

Negatively charged
Non-polar
Positively charged
Polar

Clustal Sequence alignment

Endonuclease

```
                1681                                                    1740
H.sapiens       WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
C.lupus         WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
M.musculus      WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
M.mulatta       WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
G.gallus        WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
X.tropicalis    WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
D.rerio         WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FASYKWNVSR
D.melanogaster  WQKIHESIVM DLCQVFDQEL DALEIETVQK ETIHPRKSYK MNSSCADILL FPAYKWNVSR
C.elegans       WQKIHESVVM DLCQVFDQEL DALEIQTVQK ETIHPRKSYK MNSSCADVLL FAQYKWNVSR
A.thaliana      WQKIHESVVM DLCQVLDQEL EPLEIETVQK ETIHPRKSYK MNSSCADVLL FAAHKWPMSK
S.pombe         WQKIHESVVW DLCQVLDQEL ESLQIETVQK ETIHPRKSYK MNSSCADILL LAAYKWNVSR
N.crassa        WQKIHESVVM DLCQVFDQEL EALSIESVQK ETIHPRKSYK MNSSCADIQL FASHKWNVTR
S.cerevisiae    WQKIHESIVF DICQILDGEL DVLQIESVTK ETVHPRKSYK MNSSAADITM ESVHEWEVSK
                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 MDSTTTQKFW 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 LDNTTTNKYW IDVQLRFGDY DSHDIERYTR AKFLDYSTDA QSMYPSPTGV
N.crassa        PSLLFDTKDV IESTTTNKFW IDVQLRYGDY DSHDIERYVR AKYLDYTTDS MSLYPSPTGL
S.cerevisiae    PSLLHETNDS FKGLITNKMW FDVQLRYGDY DSHDISRYVR AKFLDYTTDN VSMYPSPTGV
```

Loops sequence highly conserved:

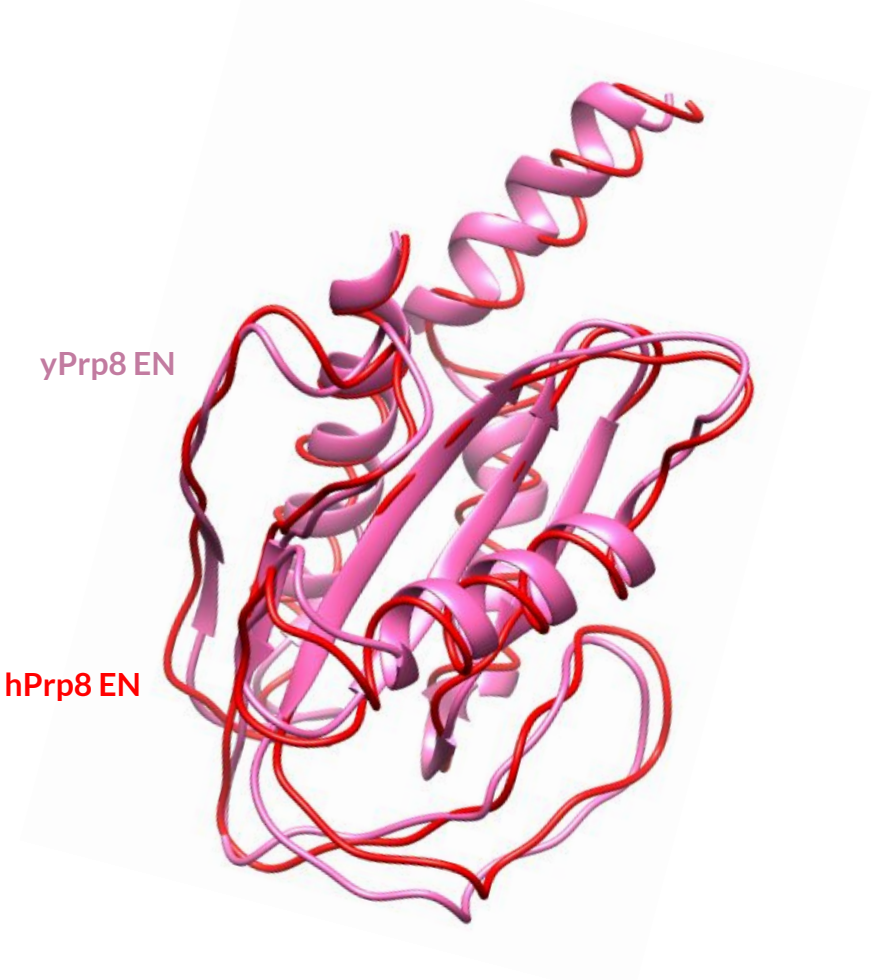
- Residues 1685-1699
- Residues 1762-1771

Clustal Sequence alignment

Endonuclease

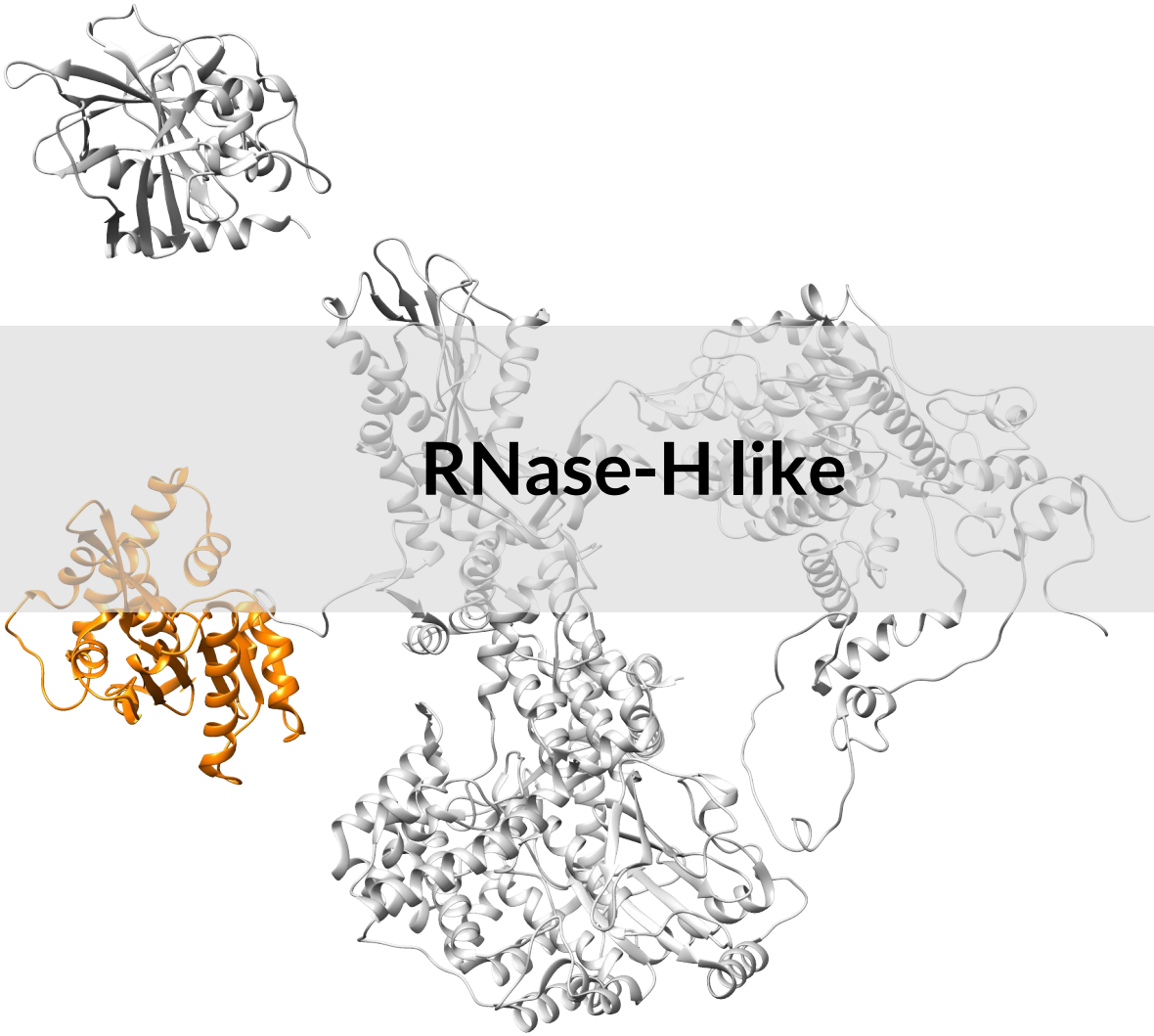
STAMP VALUES

- Sc 9.12
- RMS 1.19



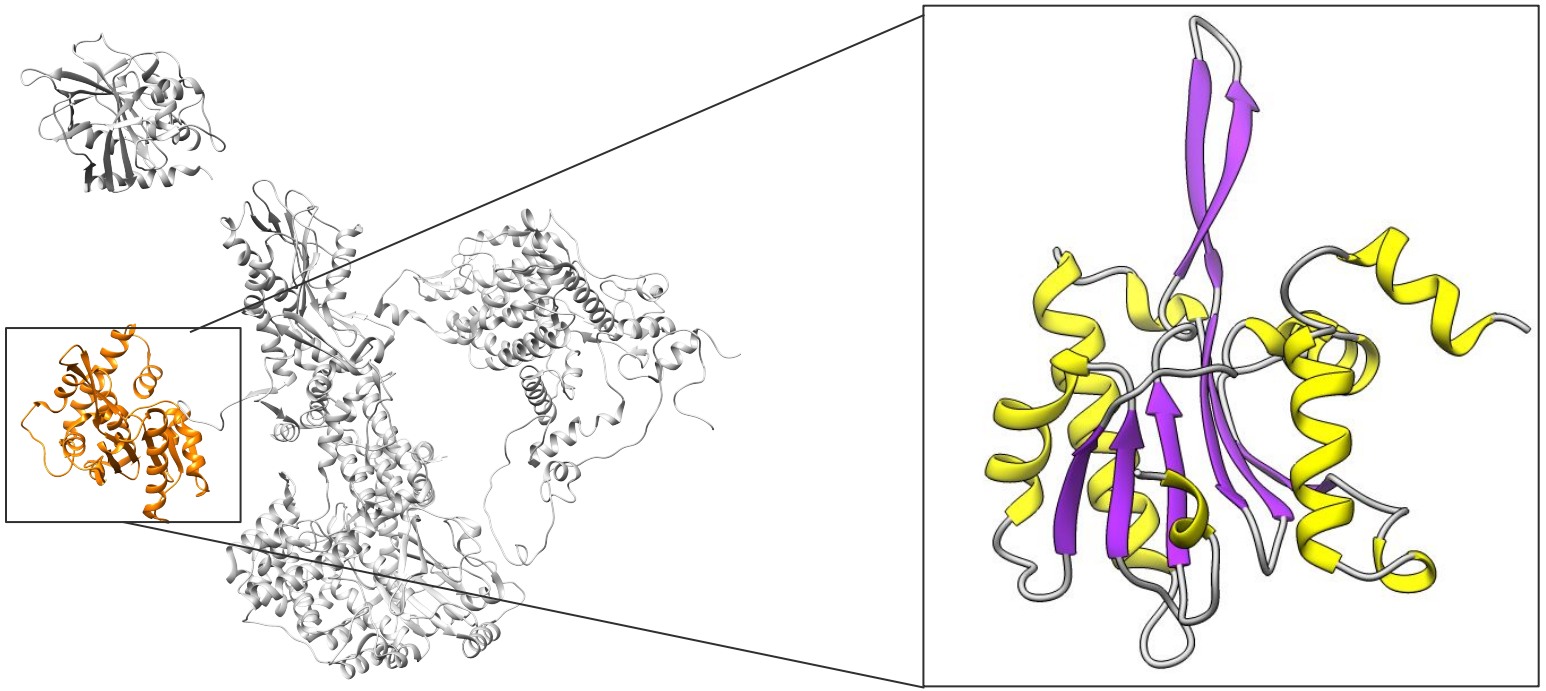
PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A



RNase-H like

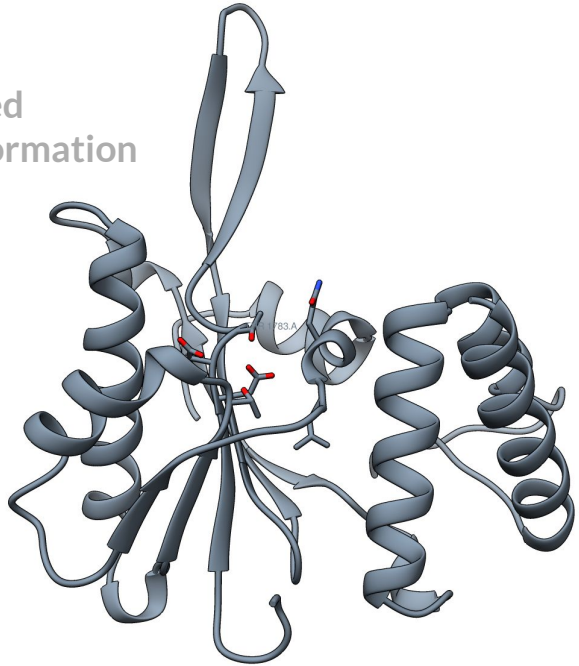
RNase-H



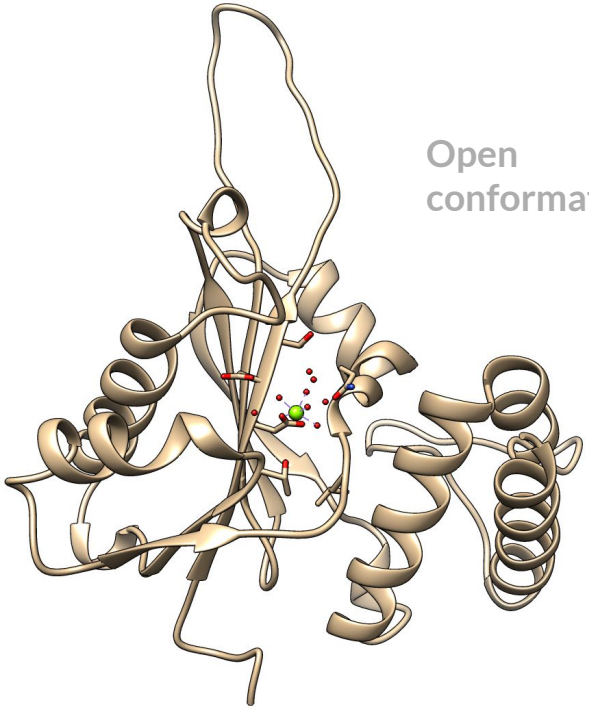
PDB code: 4I43
Resolution: 2A

RNase-H

Closed conformation

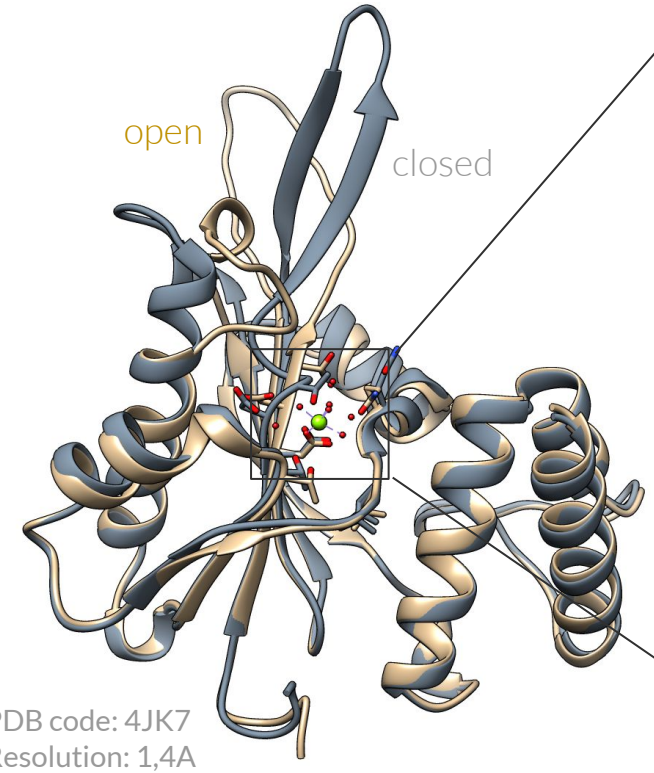


Open conformation

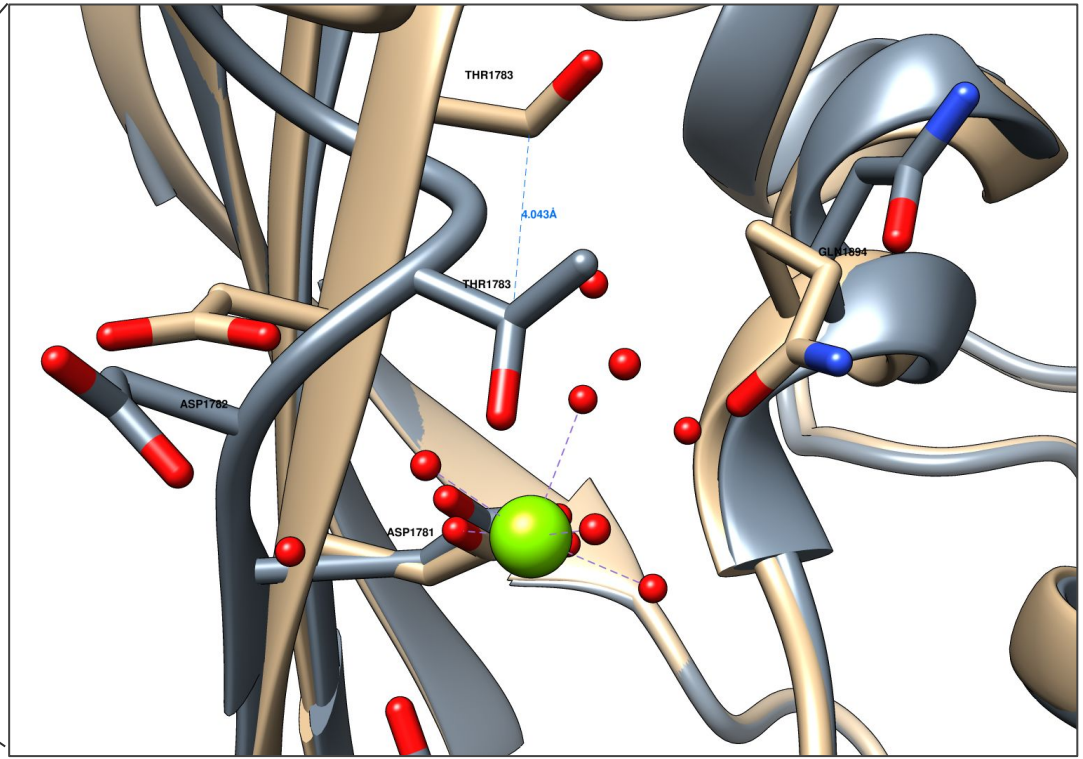


PDB code: 4JK7
Resolution: 1,4A

RNase-H

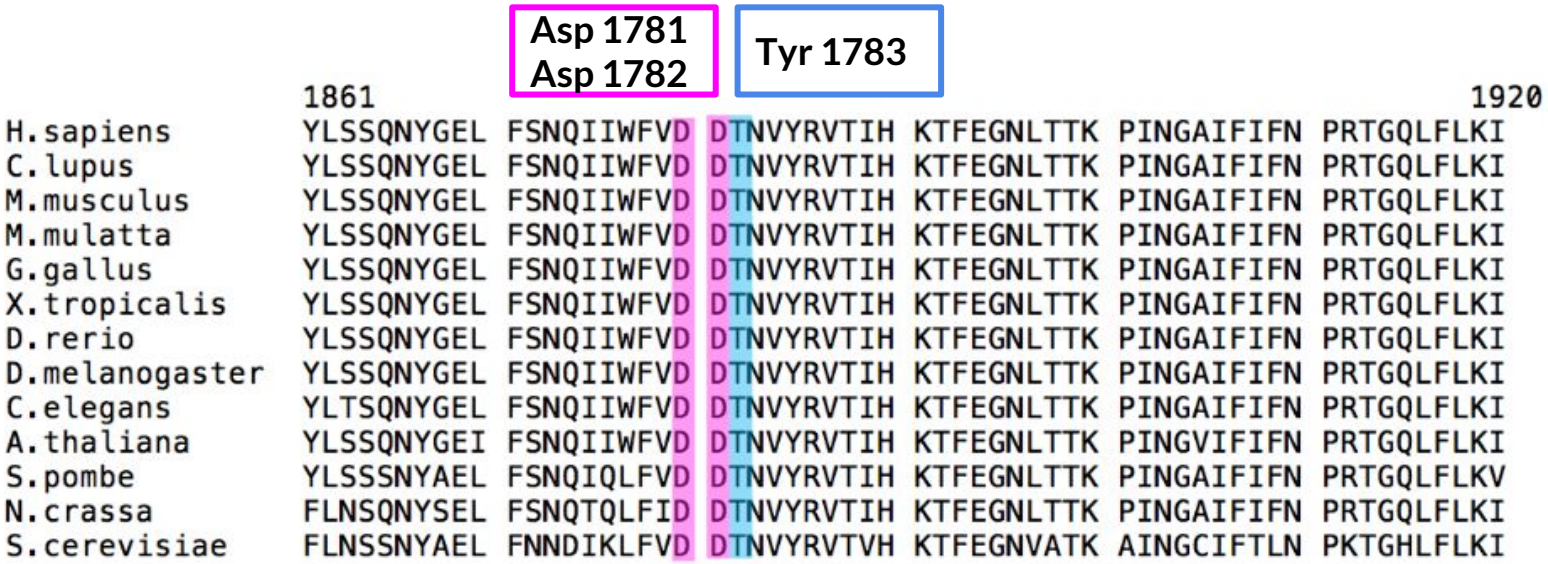


PDB code: 4JK7
Resolution: 1,4A



STAMP VALUES Sc 7.40 RMS 0.93

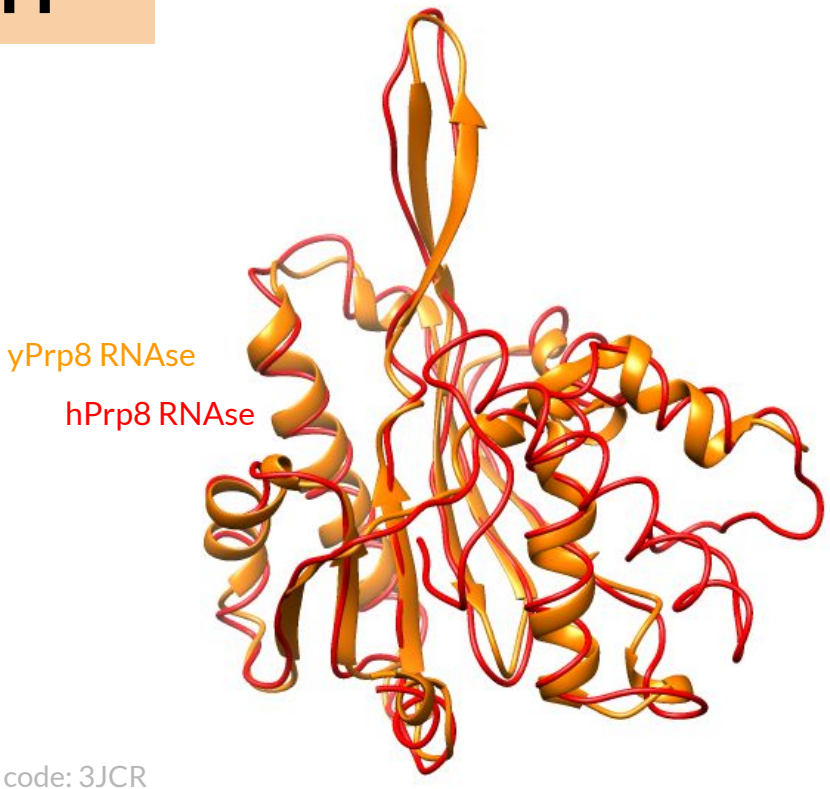
RNase-H



Negatively charged
Non-polar
Positively charged
Polar

Clustal Sequence alignment

RNase-H

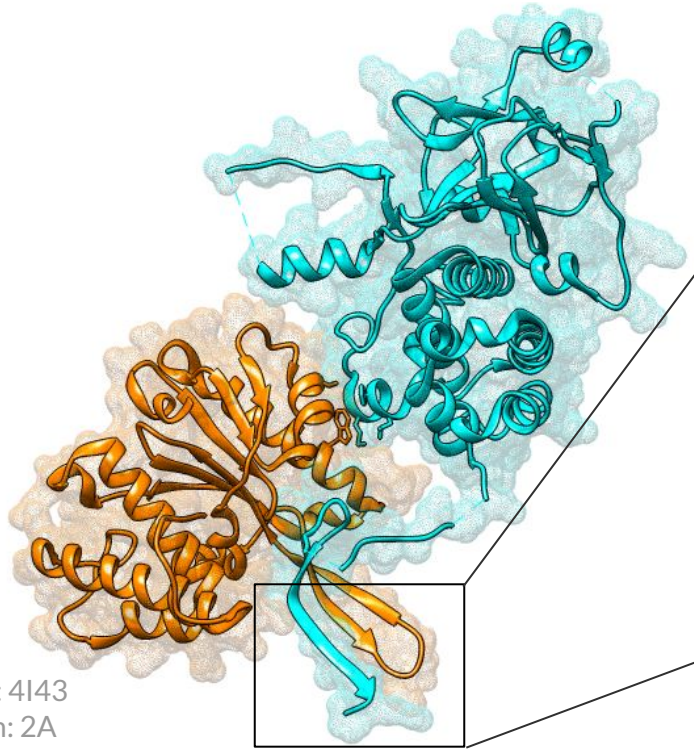


STAMP VALUES
Sc 7.11
RMS 1.24

PDB code: 3JCM
Resolution: 3.8A

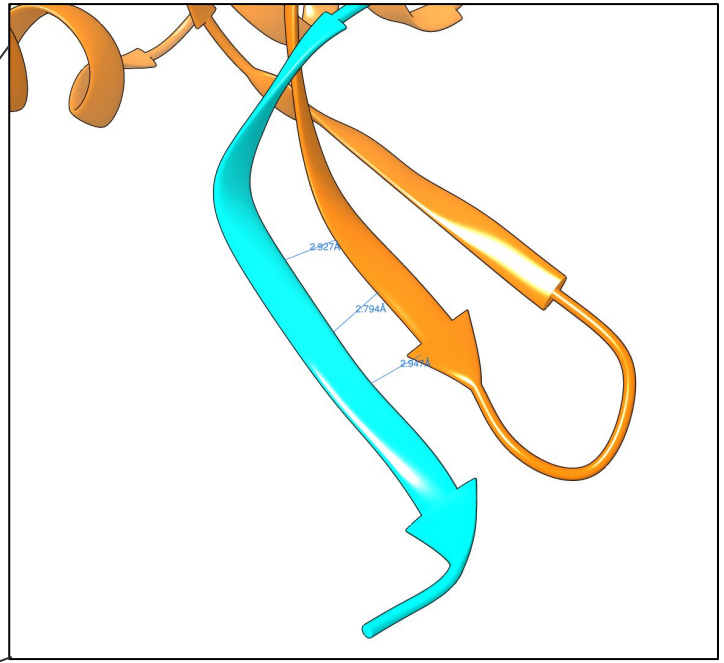
PDB code: 3JCR
Resolution: 7A

RNase-H

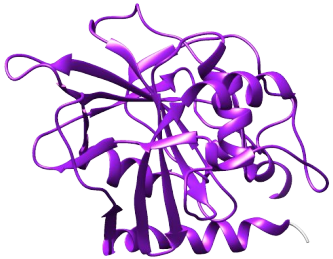


PDB code: 4I43
Resolution: 2A

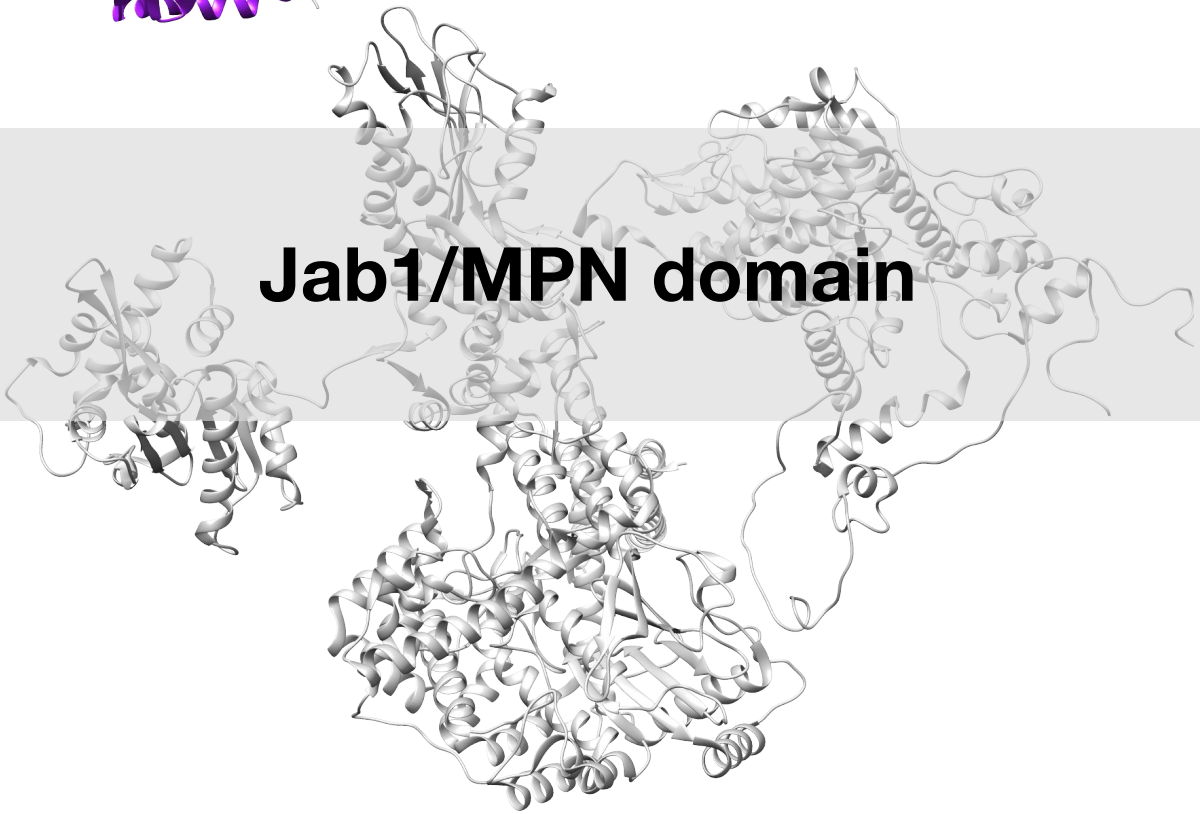
Aar2 stabilizes the β-hairpin...



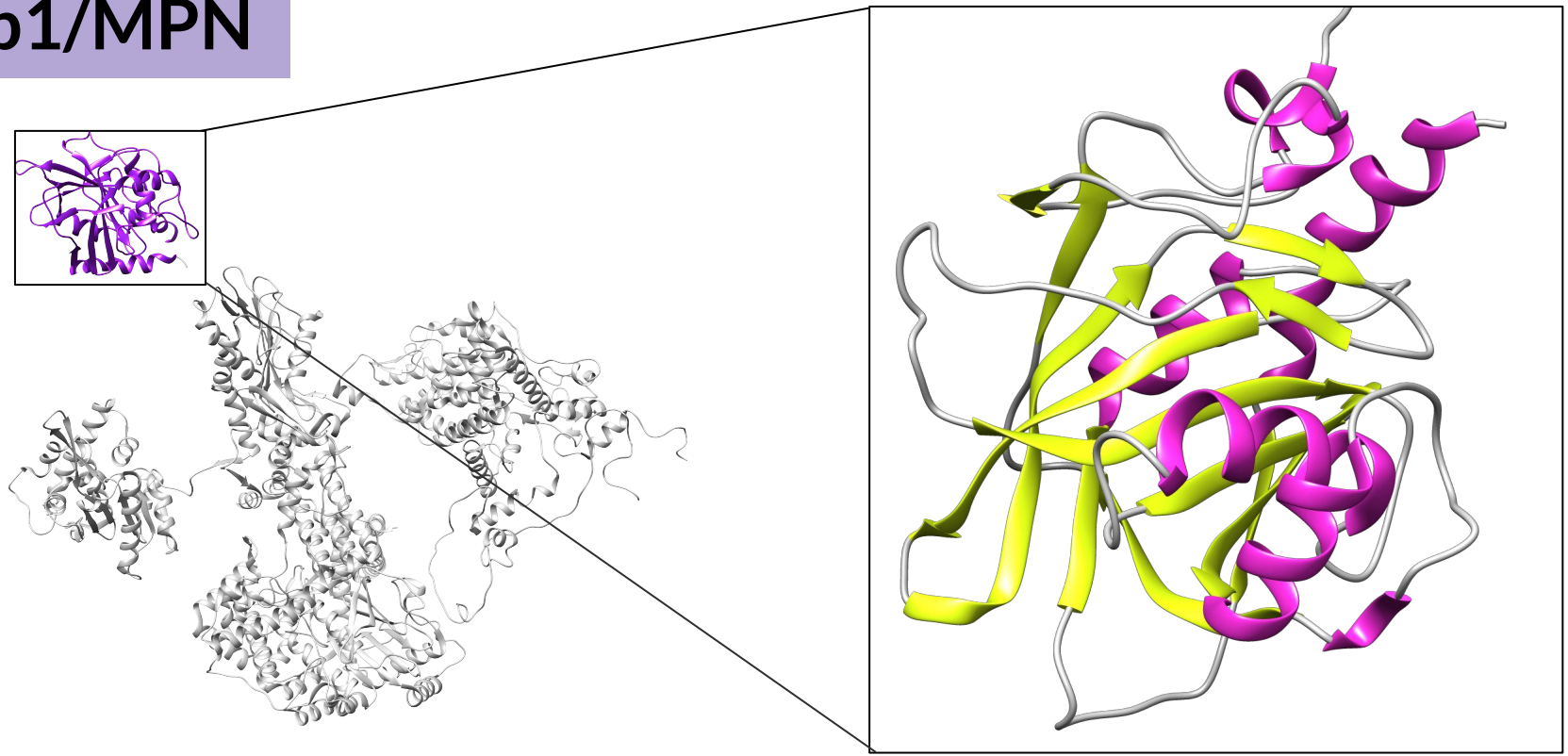
...in order to make it stay in the pre-catalytic conformation (closed)



Jab1/MPN domain



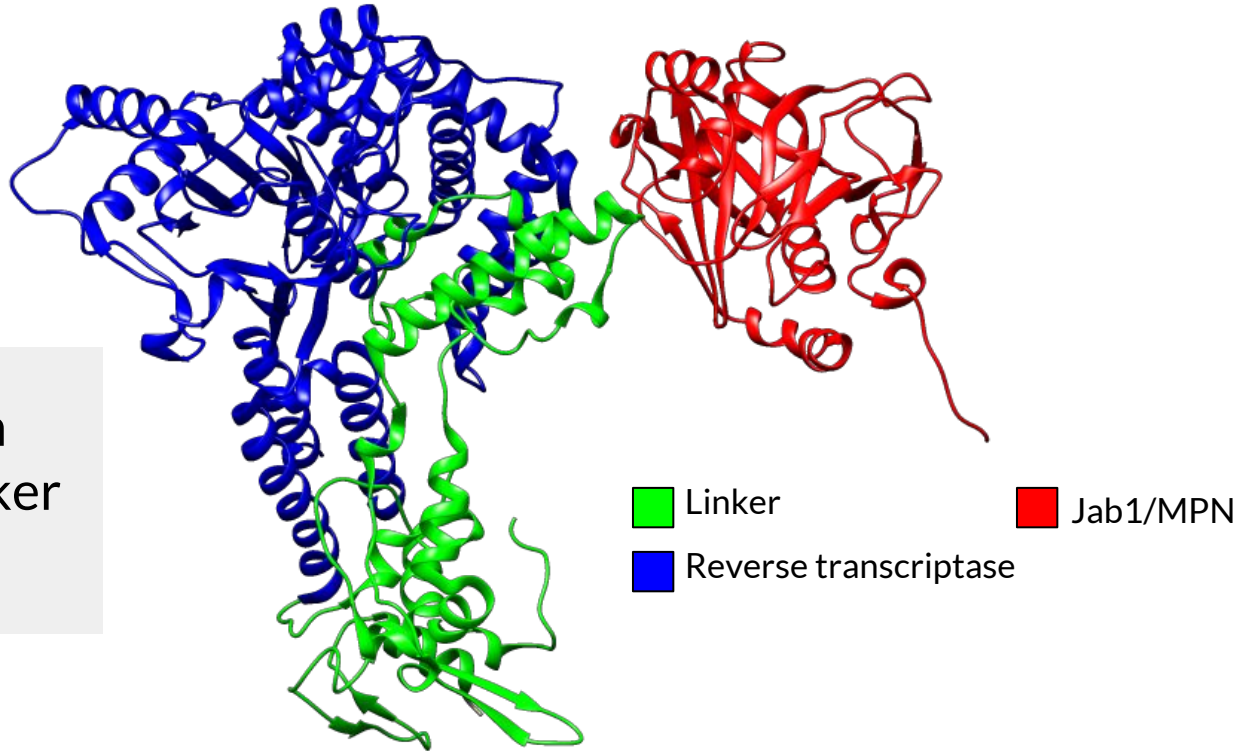
Jab1/MPN



PDB code: 4I43
Resolution: 2A

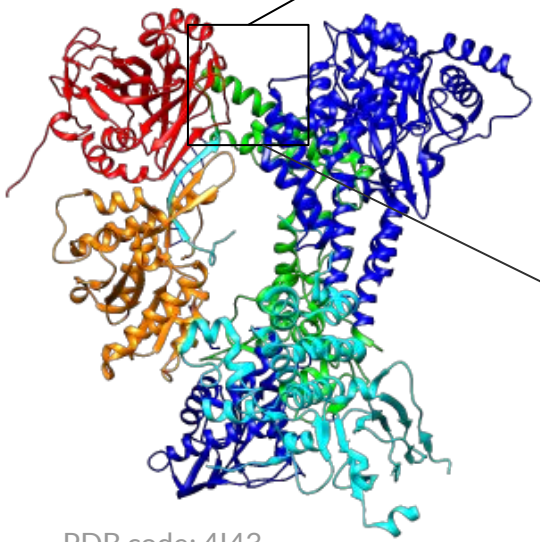
Jab1/MPN

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

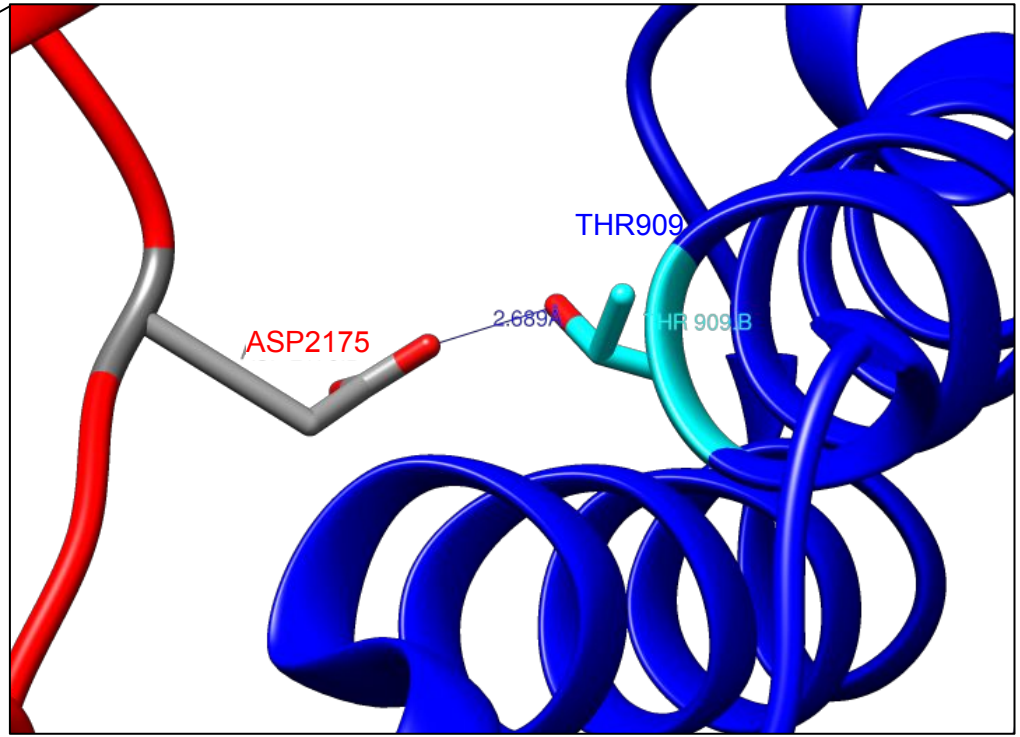


PDB code: 4I43
Resolution: 2A

Jab1/MPN



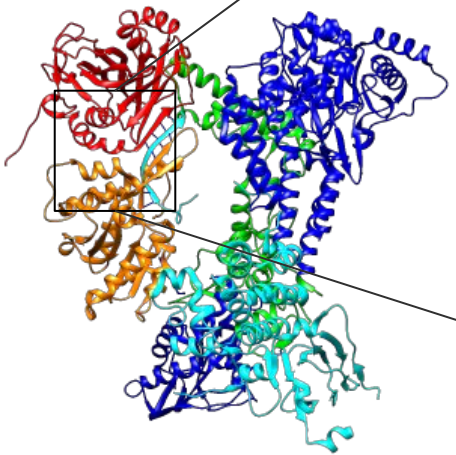
PDB code: 4I43
Resolution: 2A



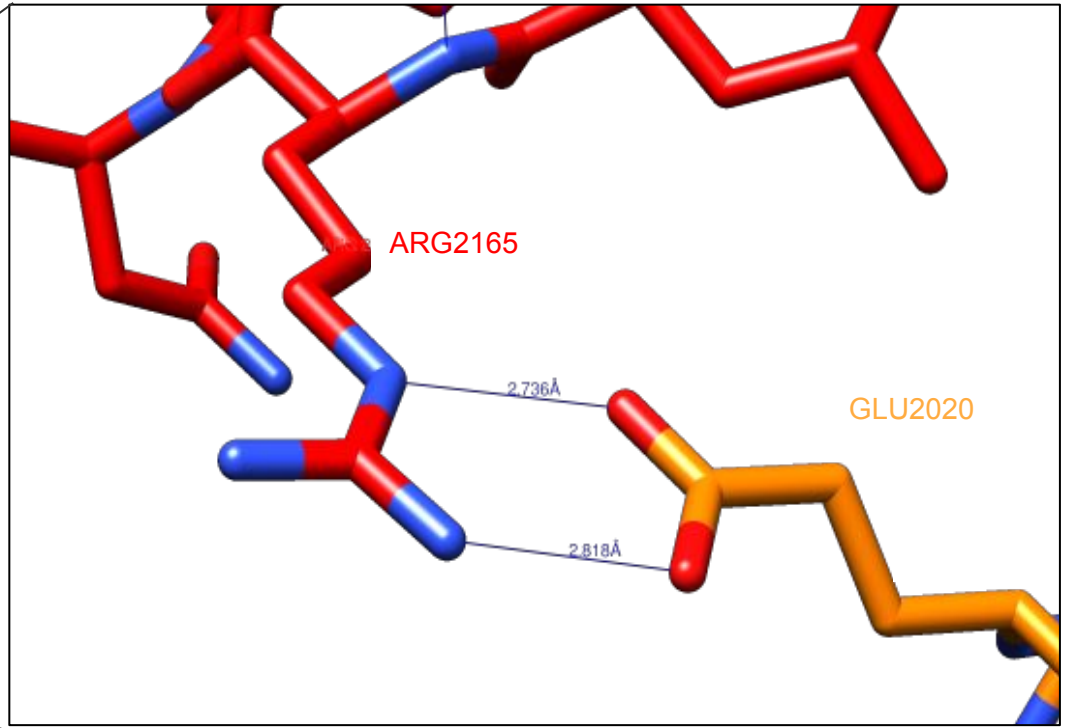
Jab1/MPN

RT

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	ILRALHVNND	RAKVILKPDK	TTITEPHHIW
M.mulatta	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	RAKVILKPDK	TTITEPHHIW
C.lupus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	RAKVILKPDK	TTITEPHHIW
M.musculus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	RAKVILKPDK	TTVTEPHHIW
G.gallus	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	RAKVILKPDK	TTITEPHHIW
D.rerio	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	RAKVILKPDK	TTITEPHHIW
D.melanogaster	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNTE	RTKIILKPDK	TTITEAHHIW
C.elegans	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRVVL	IMRGMHINPD	KTKVILKPDK	TTITEPHHIW
S.cerevisiae	ATEPQMVLFN	IYDDWLDRIS	SYTAFSRLTL	LLRALKTNEE	SAKMILLSDP	TITIKSYHLW
S.pombe	ATEPQMVLFN	LYDDWLQSVS	SYTAFSRLIL	ILRALNVNTE	KTKLILRPDK	SIITKENHVW
N.crassa	ATEPQMVLFN	LYDEWLKSIS	SYTAFSRLIL	ILRALHVNQD	KTKLILRPDK	TVITQDHHIW
A.thaliana	ATEPQMALFN	IYDDWLMTVS	SYTAFQRLIL	ILRALHVNNE	KAKMLLKPDM	SVVTEPNHIW
X.tropicalis	ATEPQMVLFN	LYDDWLKTIS	SYTAFSRLIL	ILRALHVNND	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 PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
C.lupus         AVAVYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
M.musculus      AVAVYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
M.mulatta       AVAVYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
G.gallus        AVAVYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
X.tropicalis    AVAVYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
D.rerio         AVAIYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
D.melanogaster  AVAIYTTTVH WLESRRFAPI PFPPLSYKHD TKLLILALER LKEAYSVKSR LNQSQREELG
C.elegans       AVAIYTTTVH WLESRRFSPI PFPPLSYKHD TKLLILALER LKESYSVKNR LNQSQREELA
A.thaliana      GIAIYSTTVN WLESRKFSAI PFPPLSYKHD TKLLILALER LKESYSAAVK LNQQQREELG
S.pombe         AVAIYTTFVH WLESRRFQPI PFPPLSYKHD TKLLVLALER LKEAYSVKGR LNQSQREELA
N.crassa        AVAIYTTTVH WLESRKFSPI PFPSVSYKHD TKILILALER LREAYSTKGR LNQSQREELA
S.cerevisiae    ATTIFSVMVE WLESRSFSPI PFPPLTYKND TKILVLALED LKDVYASKVR LNASEREELA
```

Negatively charged

Non-polar

Positively charged

Polar

Clustal Sequence alignment

Jab1/MPN

Jab1/MPN region

Arg 2165

Asp 2175

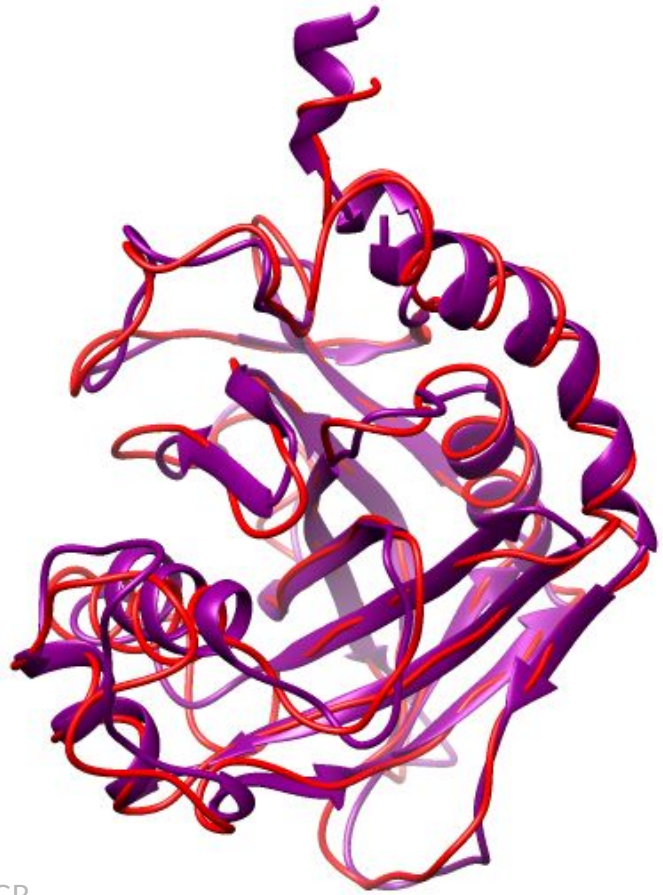
	2161					2220
H.sapiens	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH	LRTNHIYVSS	DDIKETG.YT	YILPKNVLKK
C.lupus	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH	LRTNHIYVSS	DDIKETG.YT	YILPKNVLKK
M.musculus	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH	LRTNHIYVSS	DDIKETG.YT	YILPKNVLKK
M.mulatta	ITSTTSNYET	QTFSSKTEWR	VRAISAANLH	LRTNHIYVSS	DDIKETG.YT	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	ITATTSNYET	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	VDSDLDD.VT	YVMPKNILKR
S.cerevisiae	VVVASADYES	QTFSSKNEWR	KSAIANTLLY	LRLKNIYVSA	DDFVEEQ.NV	YVLPKNLLKK

Negatively charged | Non-polar

Positively charged | Polar

Clustal Sequence alignment

Jab1/MPN



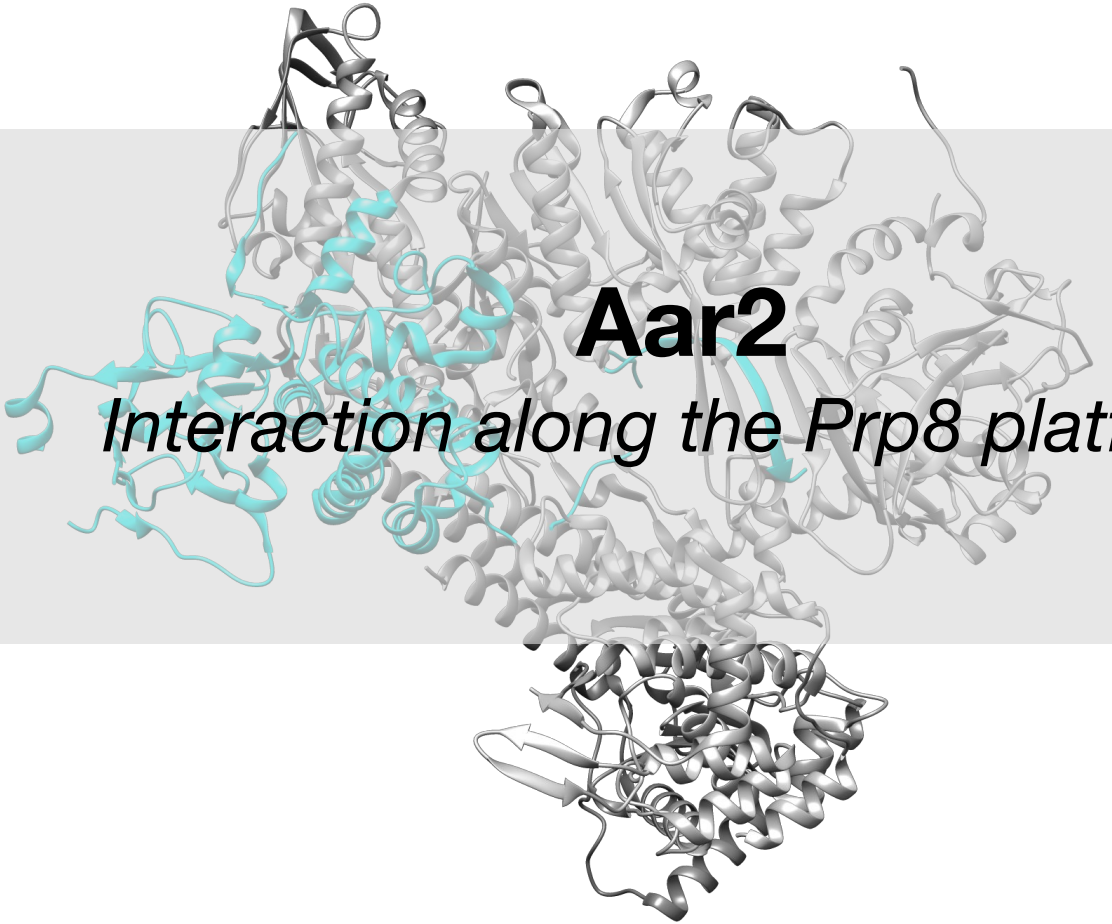
yPrp8 Jab1

hPrp8 Jab1

STAMP VALUES
Sc 8,02
RMS 1.29

PDB code: 3JCM
Resolution: 3.8A

PDB code: 3JCR
Resolution: 7A

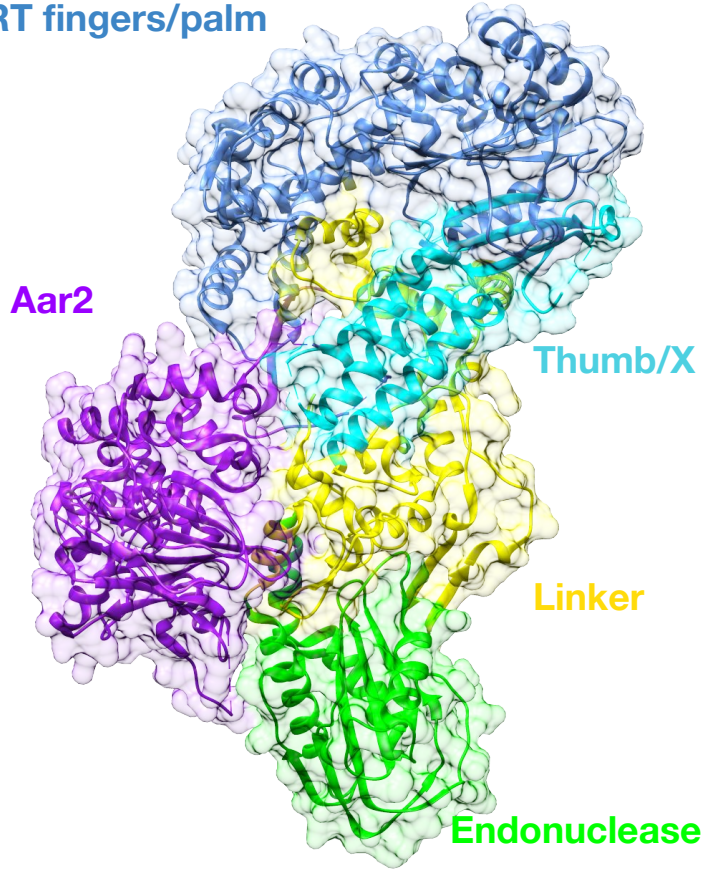


Aar2

Interaction along the Prp8 platform

Aar2

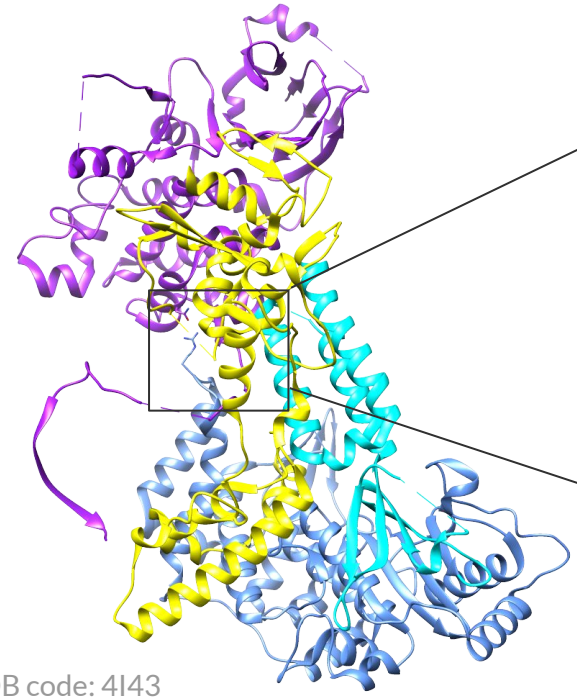
90°

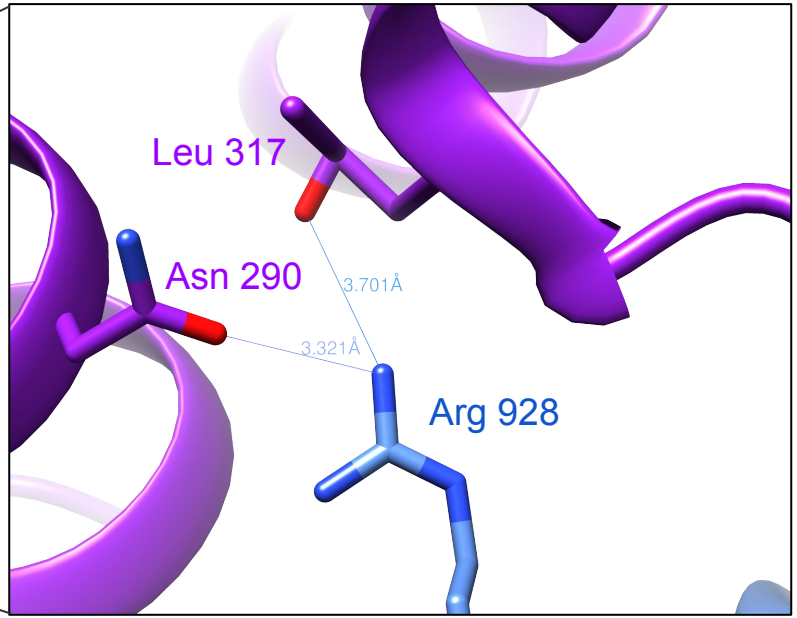


Aar2
RT
Jab1/MPN
Linker
Endonuclease

Aar2

Aar2 - RT





PDB code: 4I43
Resolution: 2A

RT fingers/palm Aar2

Aar2

Aar2 - RT

RT region

Arg 928

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

Negatively charged

Non-polar

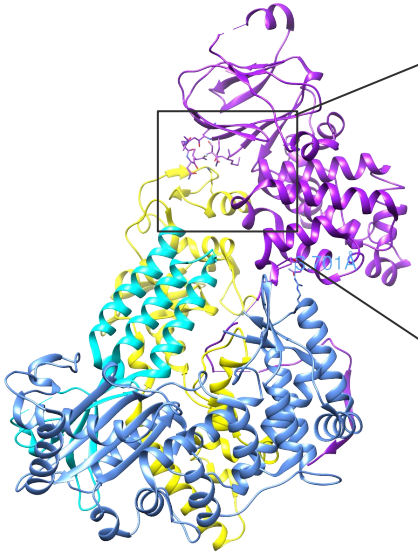
Positively charged

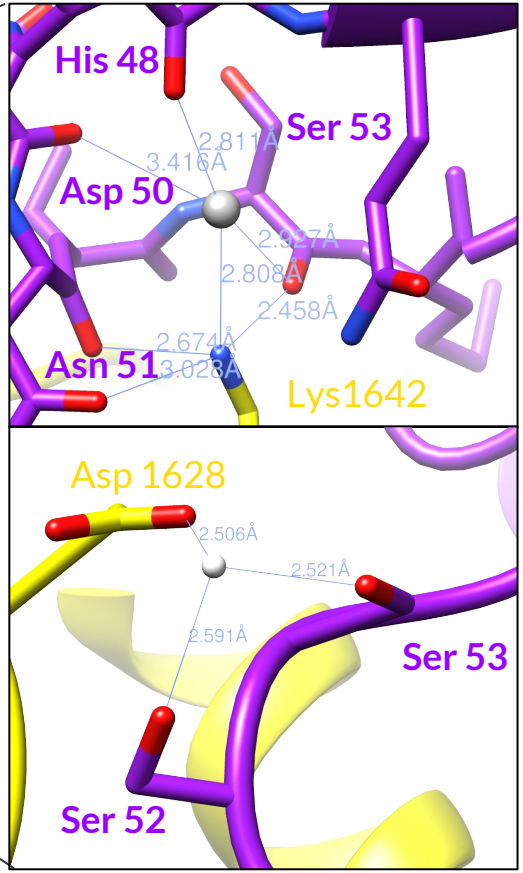
Polar

Clustal Sequence alignment

Aar2

Aar2 - Linker





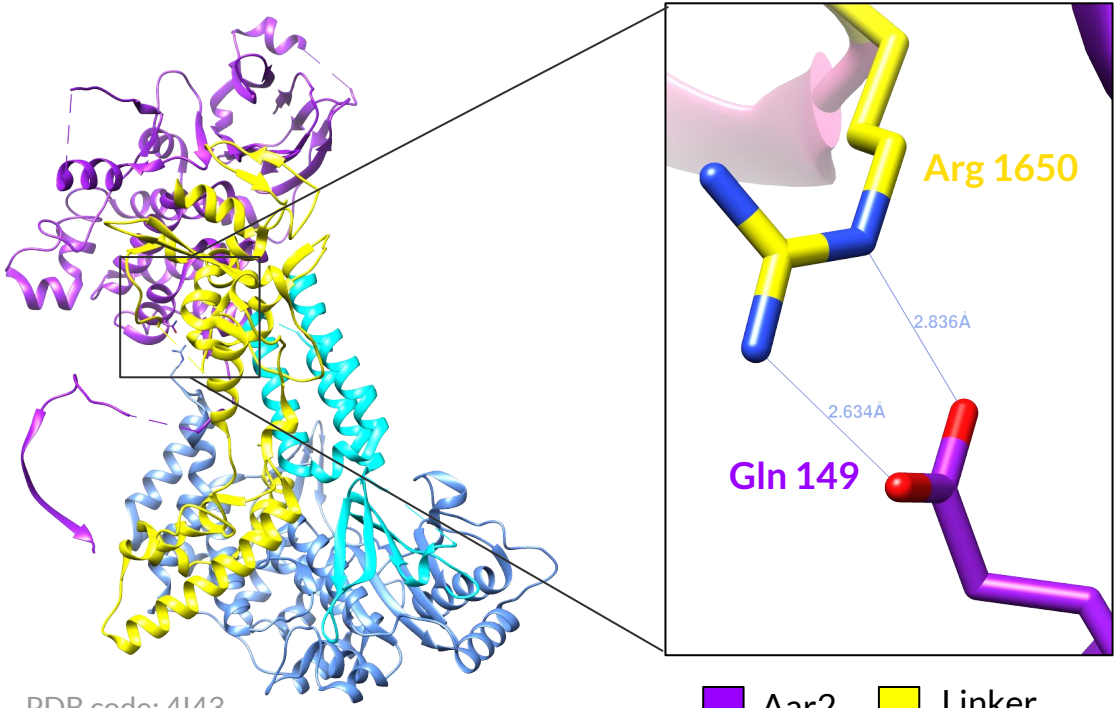
Aar2

Linker

PDB code: 4I43
Resolution: 2A

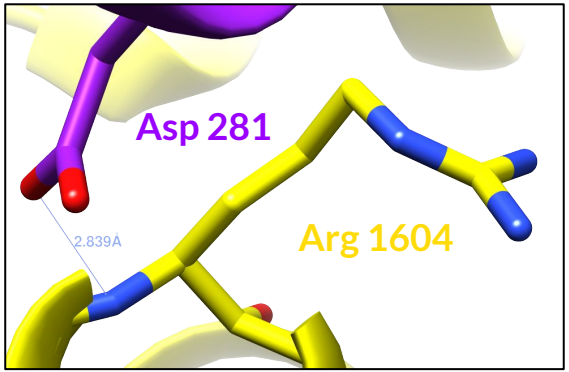
Aar2

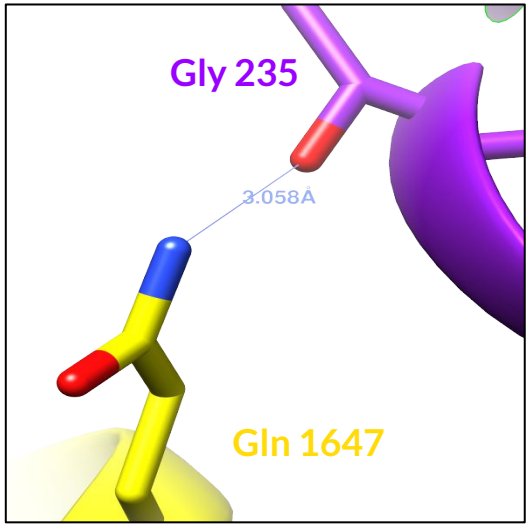
Aar2 - Linker



PDB code: 4I43
Resolution: 2A

Aar2 Linker





Aar2

Aar2 - Linker

Linker region

Arg 1604

	1561					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	LFKGTGFNSW	EGLFWEKASG	FEDSMQFKKL	THAQRTGLSQ	IPNRRFTLWW	SPTINRANVY
S.pombe	MFKATGFPSW	EGLFWEKASG	FEESMKFKKL	TNAQRSGLNQ	IPNRRFTLWW	SPTINRANVY
N.crassa	LFKATGFPSW	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

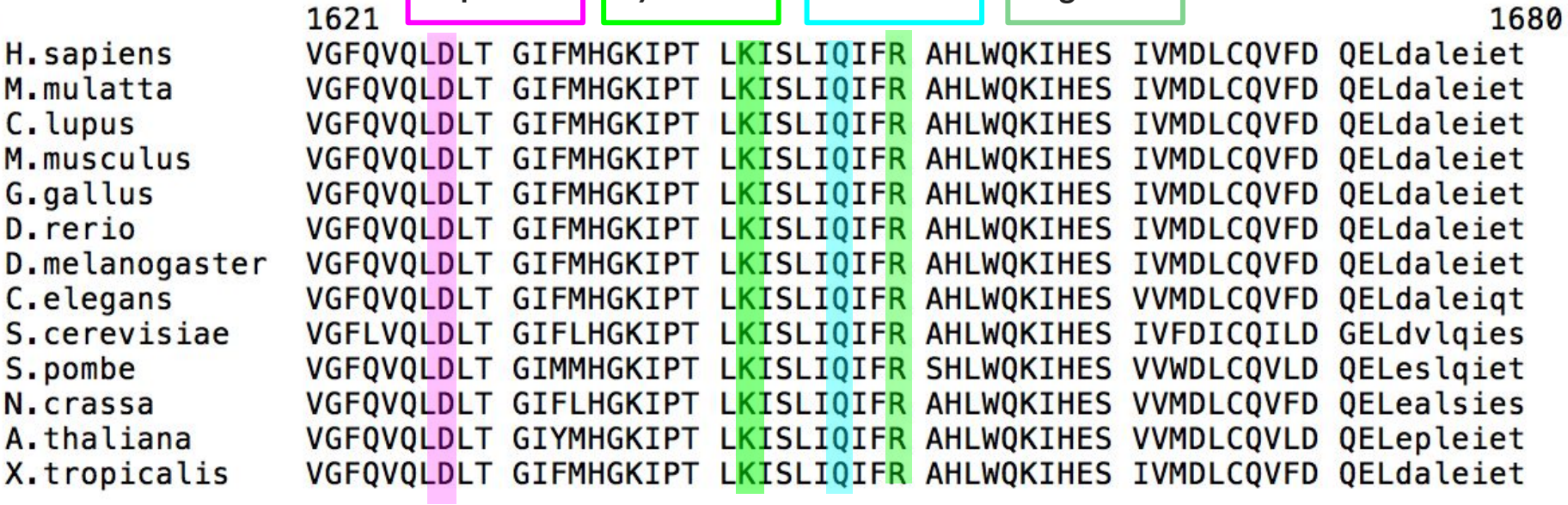
Sequence alignment based on PFAM domains

Aar2

Aar2 - Linker

Asp 1628 | Lys 1690 | Gln 1647 | Arg 1650

Linker region

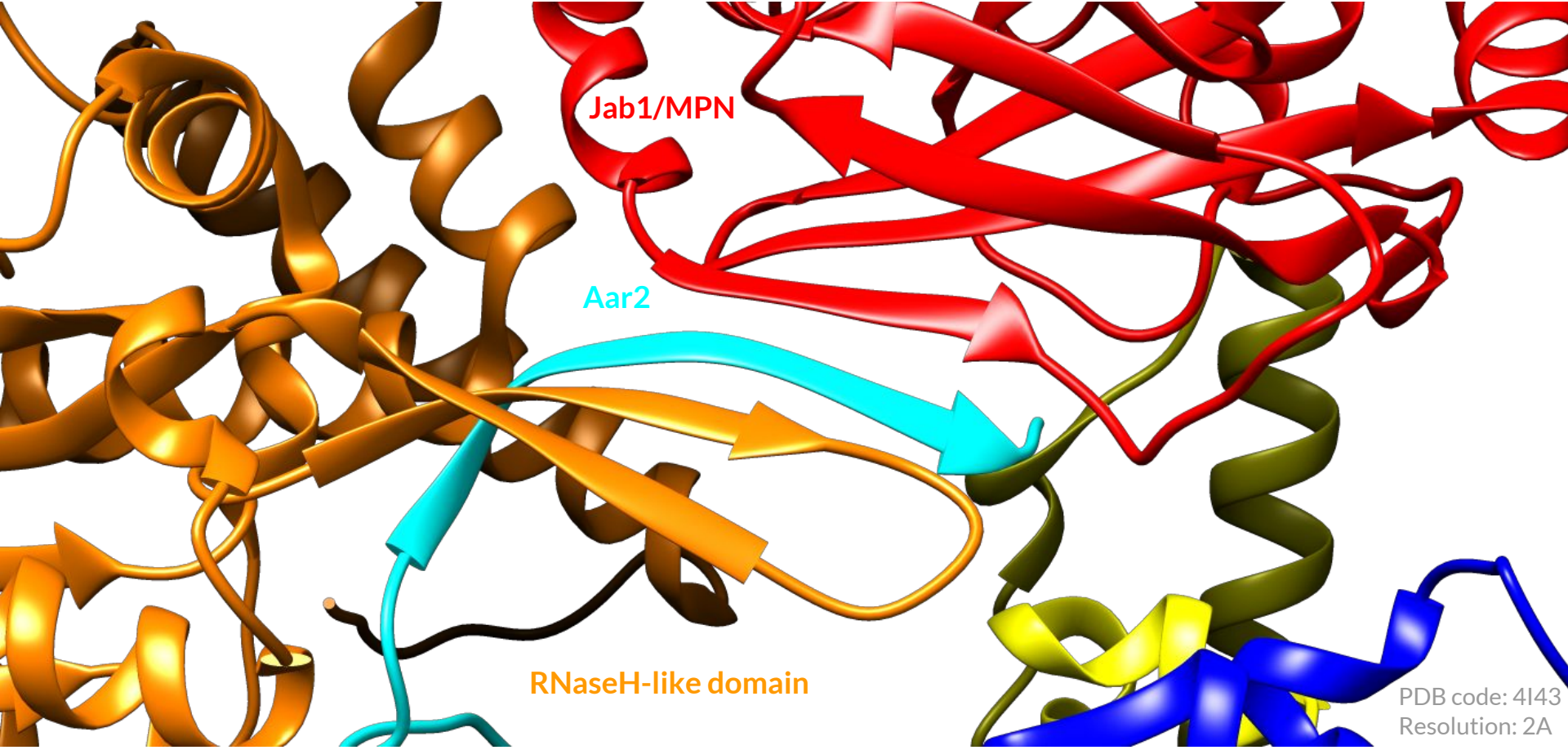


Negatively charged | Non-polar

Positively charged | Polar

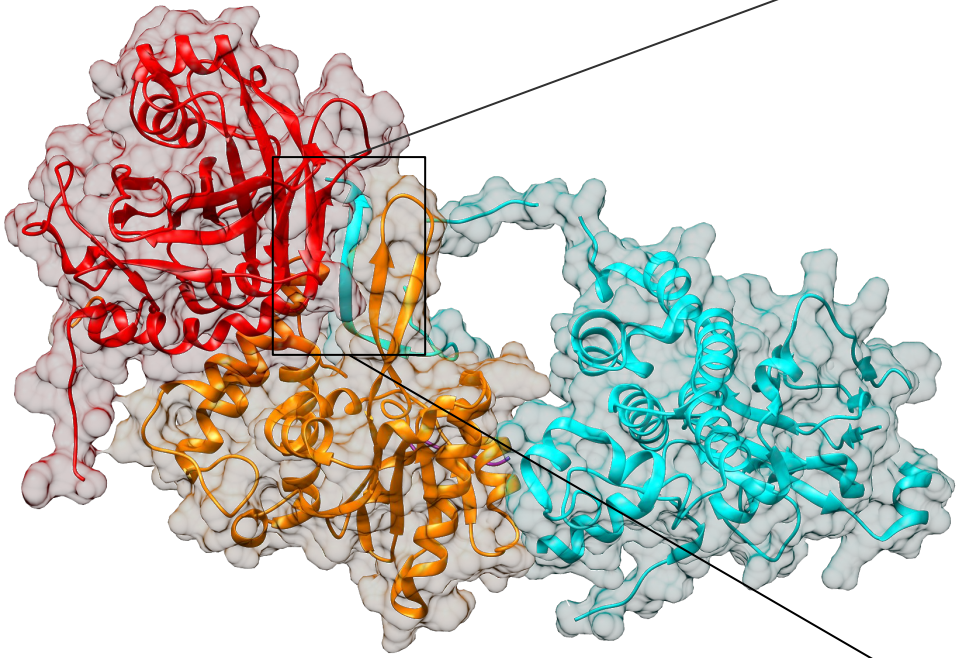
Sequence alignment based on PFAM domains

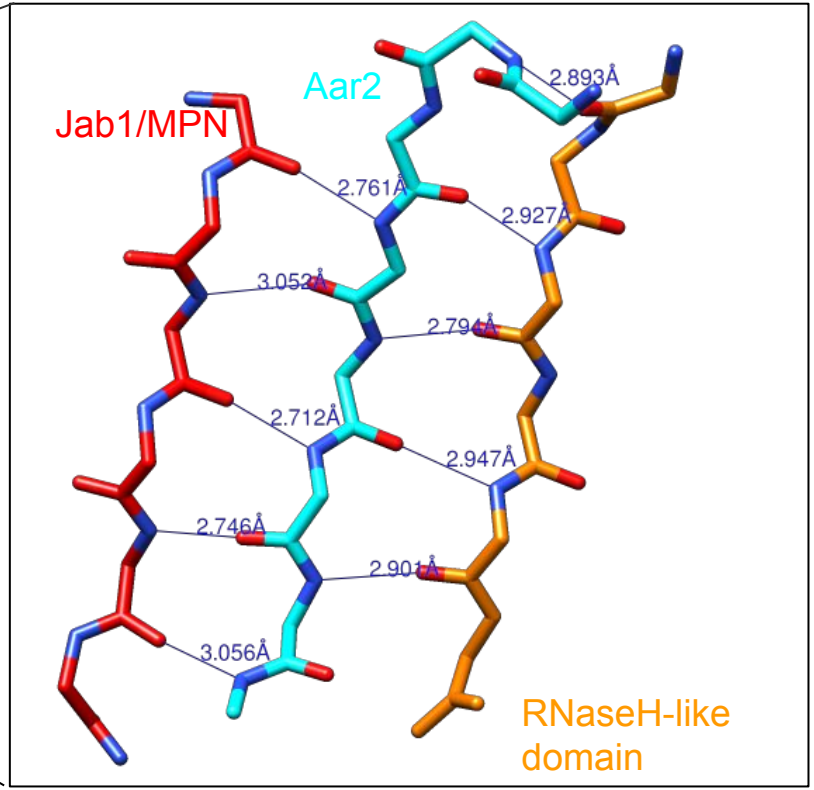
Aar2


Jab1/MPN
Aar2
RNaseH-like domain
PDB code: 4I43
Resolution: 2A

Aar2

Aar2 - Jab1/RNase





PDB code: 4I43
Resolution: 2A

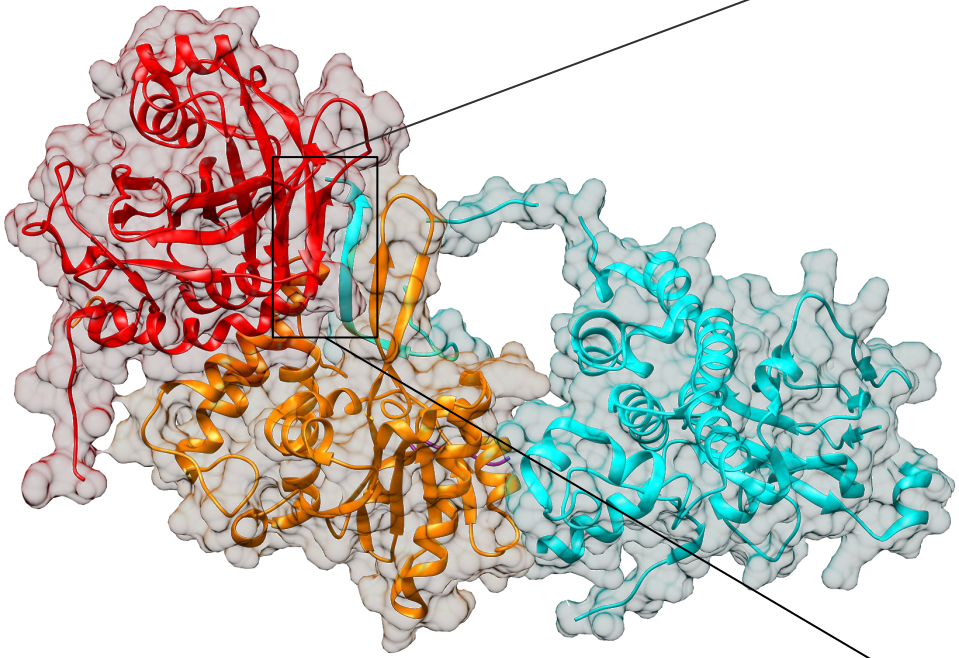
RNaseH-like

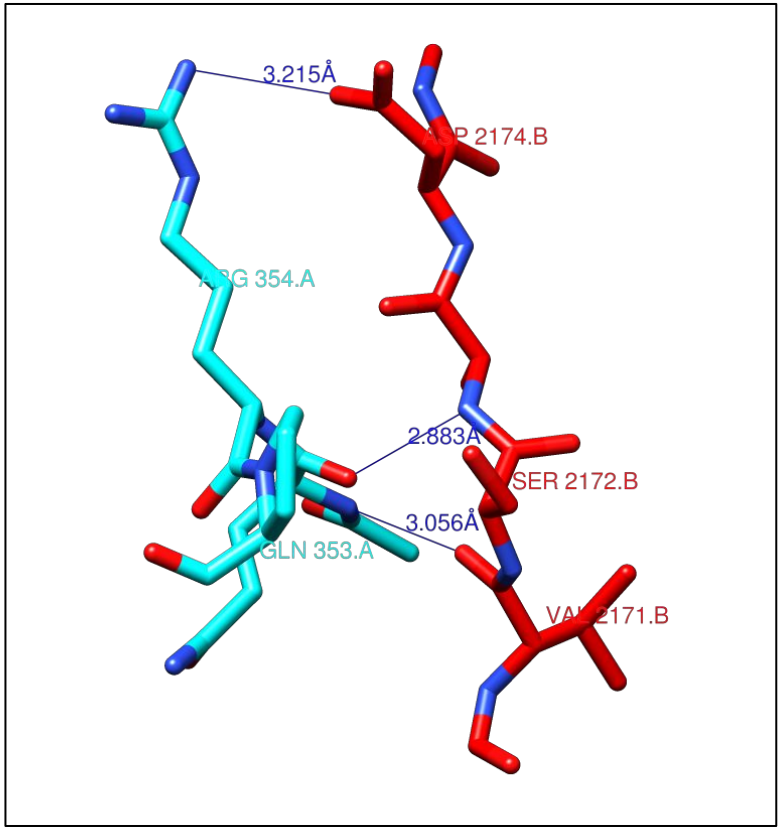
Aar2

Jab1/MPN

Aar2

Aar2 - Jab1/RNase





PDB code: 4I43
Resolution: 2A

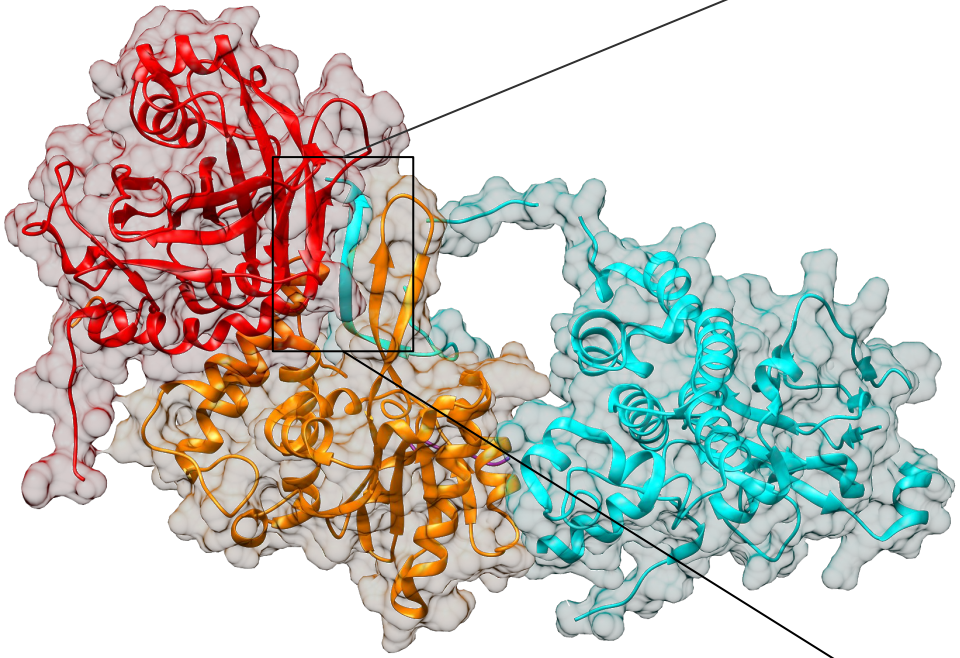
RNaseH-like

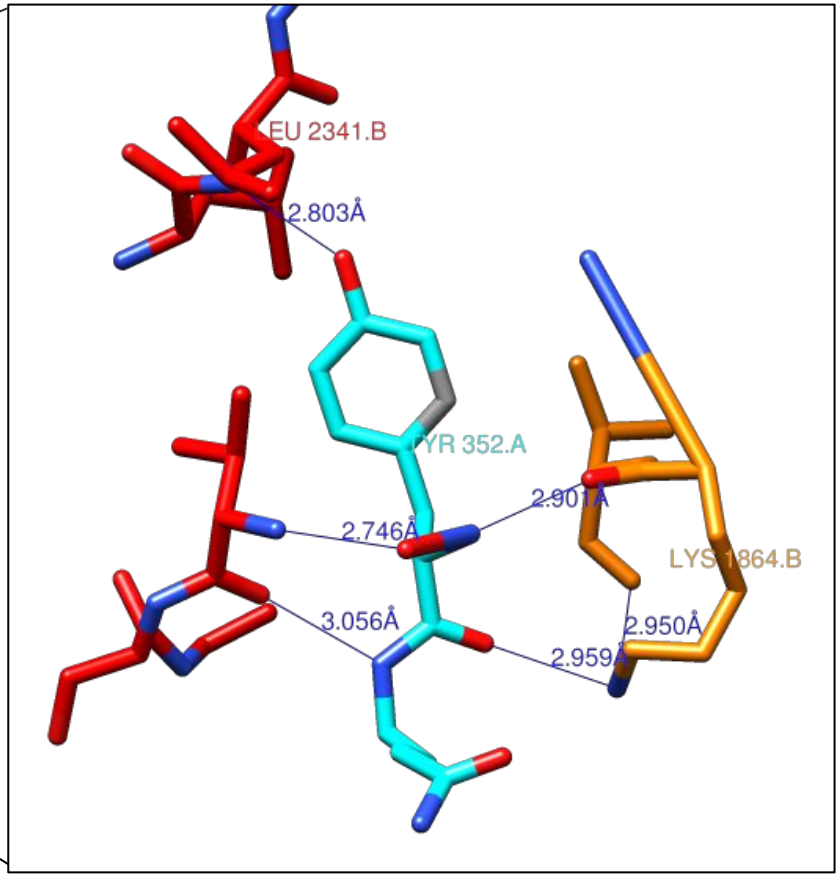
Aar2

Jab1/MPN

Aar2

Aar2 - Jab1/RNase





PDB code: 4I43
Resolution: 2A

RNaseH-like

Jab1/MPN

Aar2

Aar2

Leu 2341

Jab1/MPN region

	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	PSHFLNFASL
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	LSEKFRGFFL	VPDGGKWNYS	FMGSAFGglE	KKPVHVKLDT	PLPFYSDQHR	PIHFSSFNEL
A.thaliana	LSDRFFGFYM	VPENGPWNYN	FMGANHT..V	SINYSLTLGT	PKEYYHQVHR	PTHFLQFSKM
X.tropicalis	LSDRFLGFFM	VPAQASWNYN	FMGVRHD..P	NMKYELQLAN	PKEFYHEVHR	PSHFLNFALL

Negatively charged | Non-polar

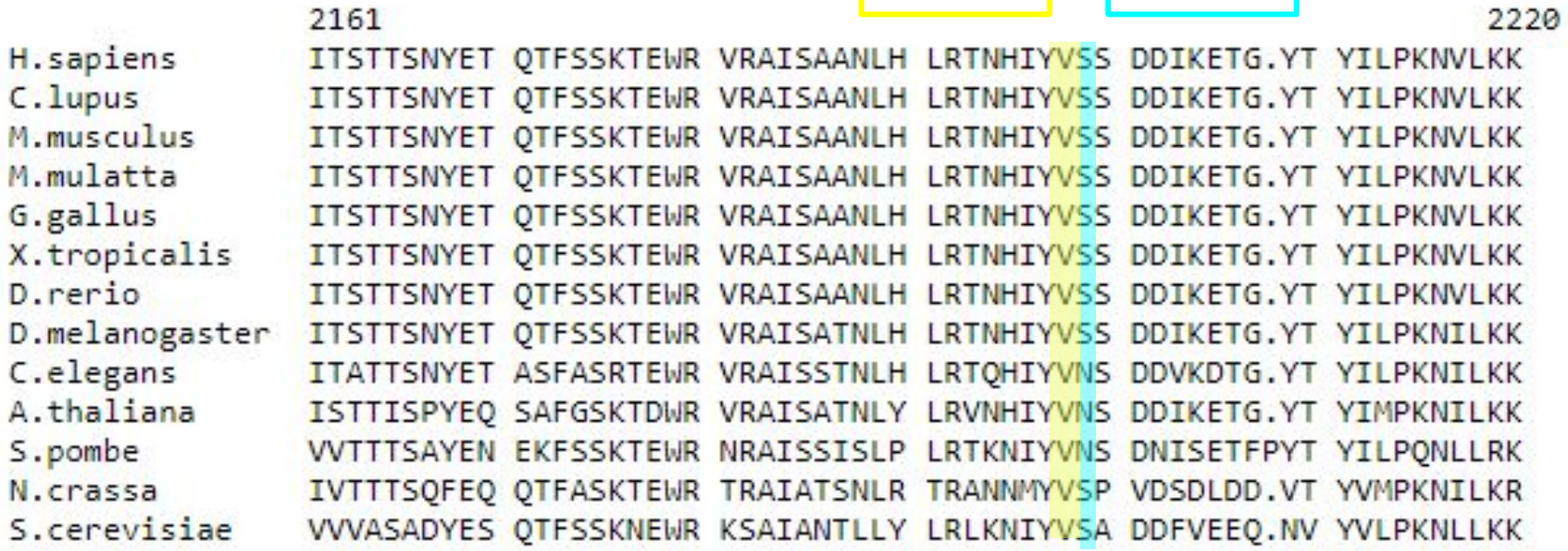
Positively charged | Polar

Sequence alignment based on PFAM domains

Aar2

Jab1/MPN region

Val 2171 **Ser 2172**



Negatively charged

Non-polar

Positively charged

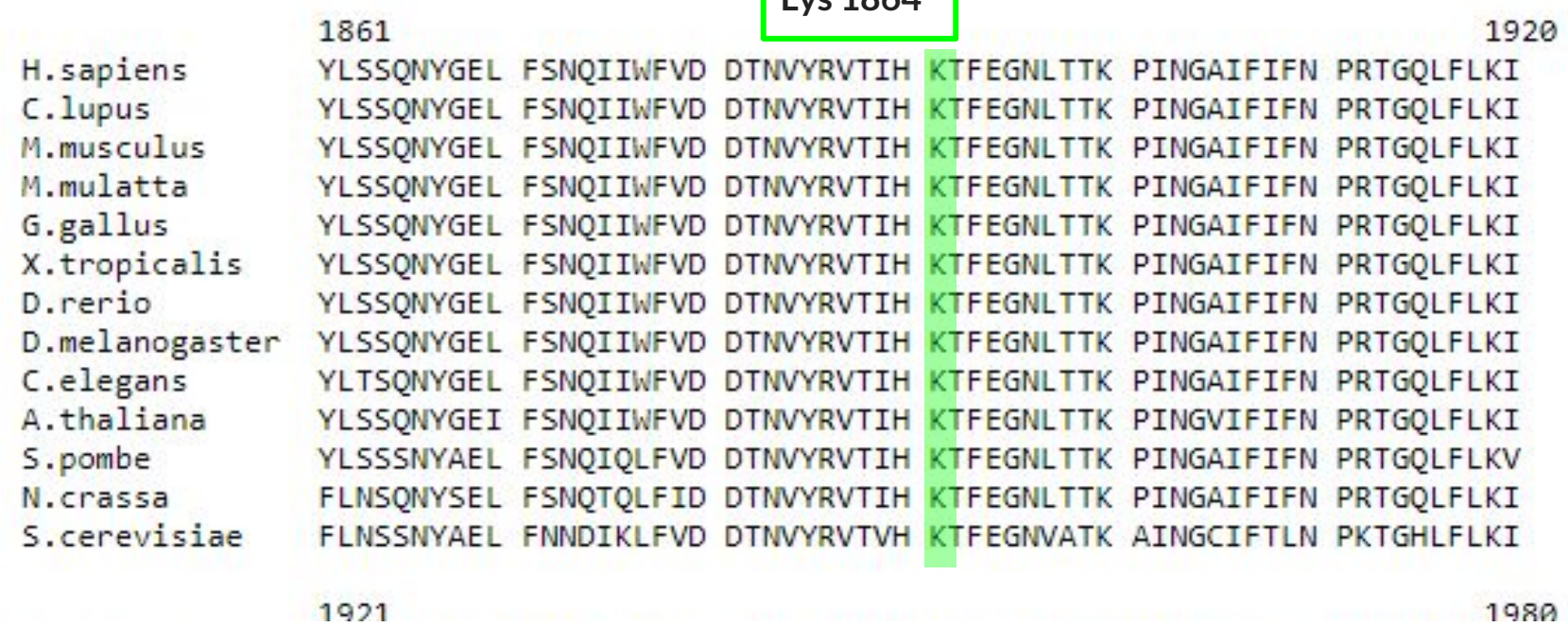
Polar

Clustal Sequence alignment

Aar2

Lys 1864

Jab1/MPN region



Negatively charged

Non-polar

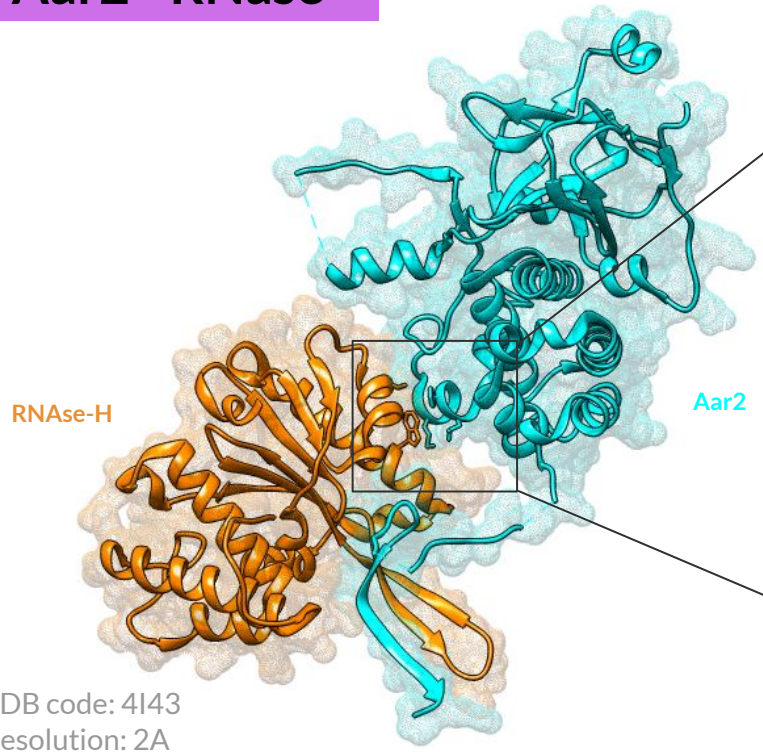
Positively charged

Polar

Clustal Sequence alignment

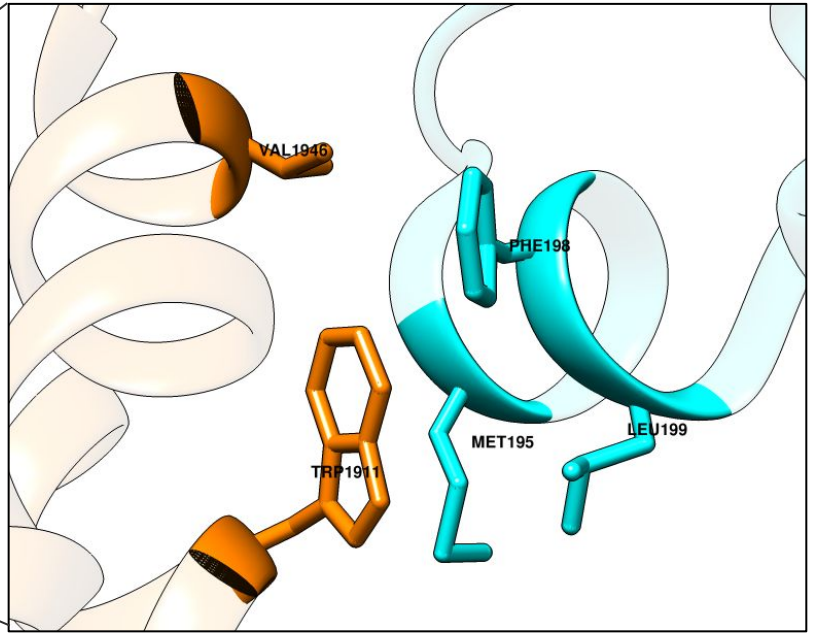
Aar2

Aar2 - RNase



PDB code: 4I43
Resolution: 2A

Interaction of the hydrophobic residues

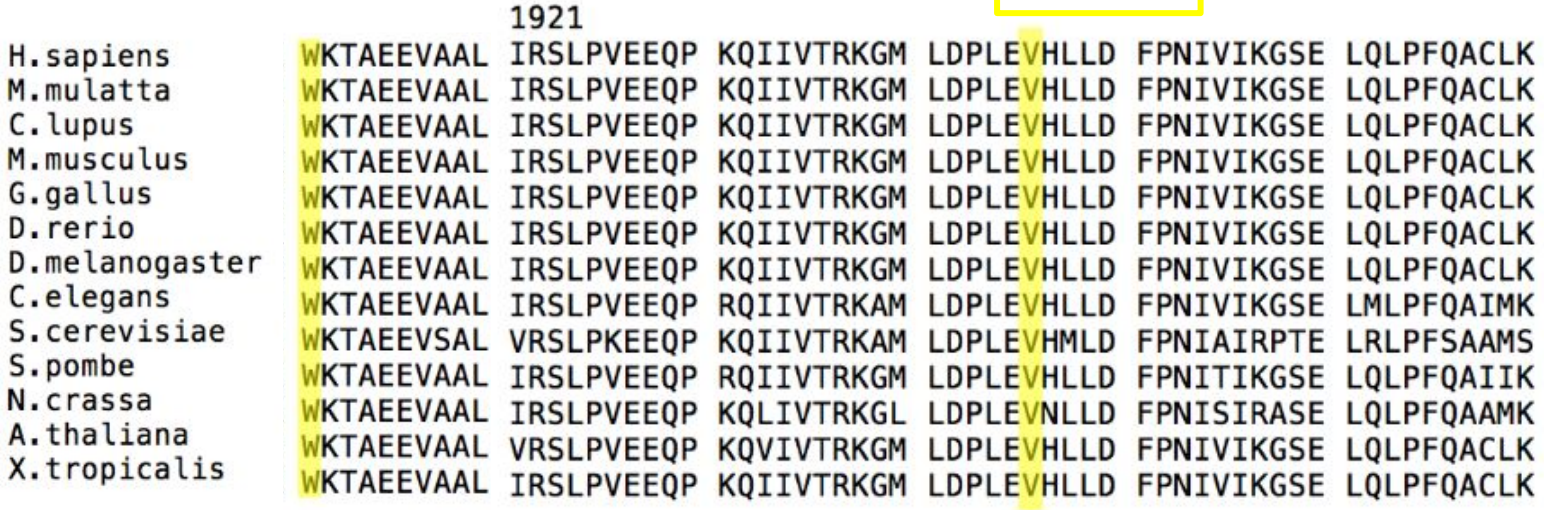


Aar2

Trp1911

Val1946

RNaseH-like region



Negatively charged

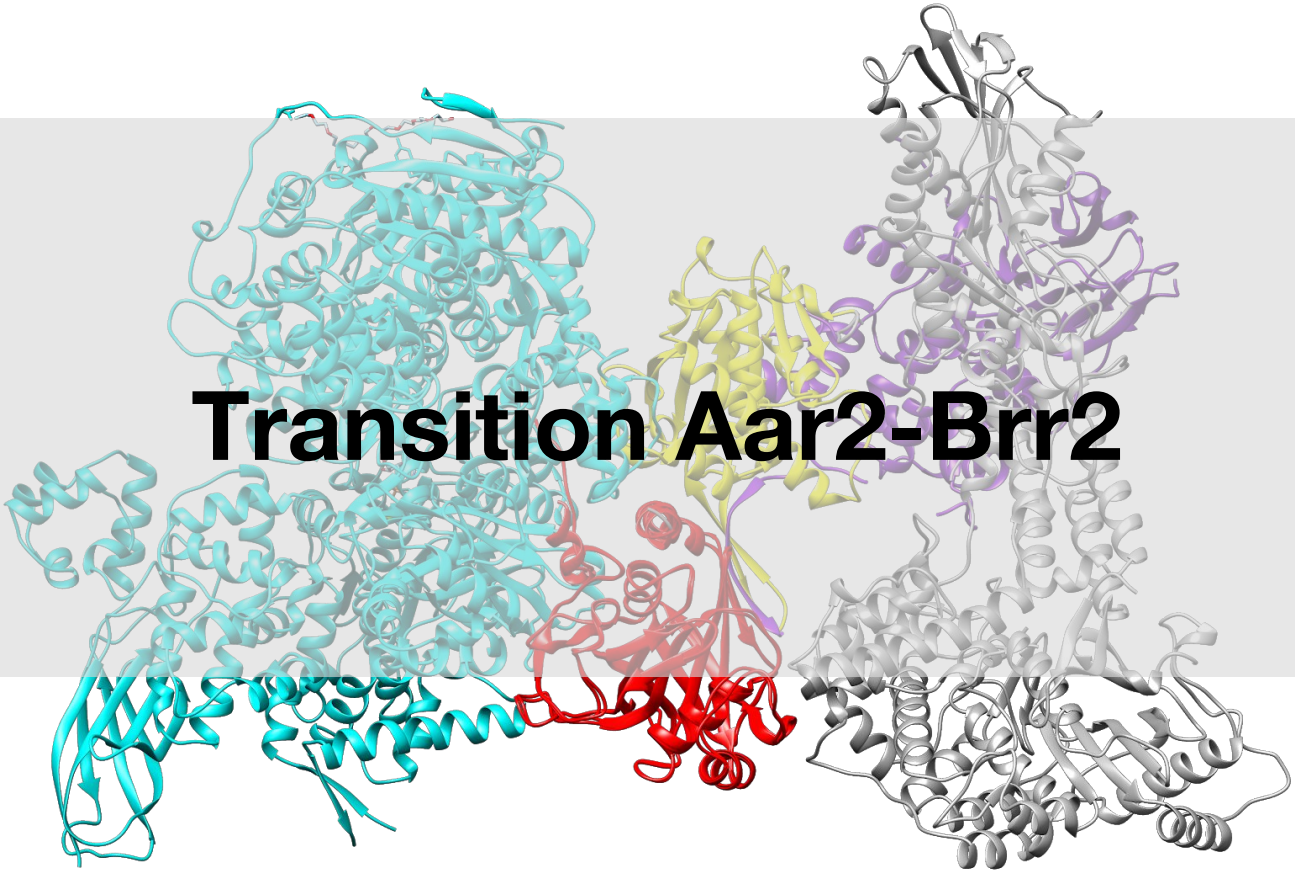
Non-polar

Positively charged

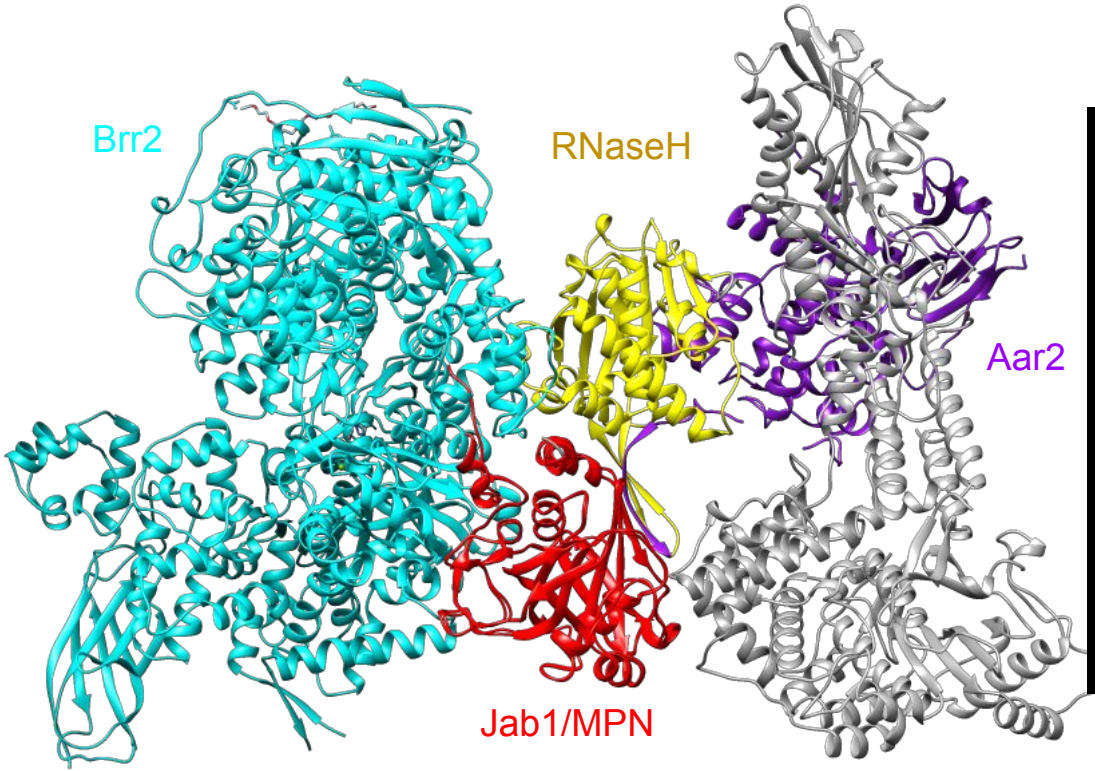
Polar

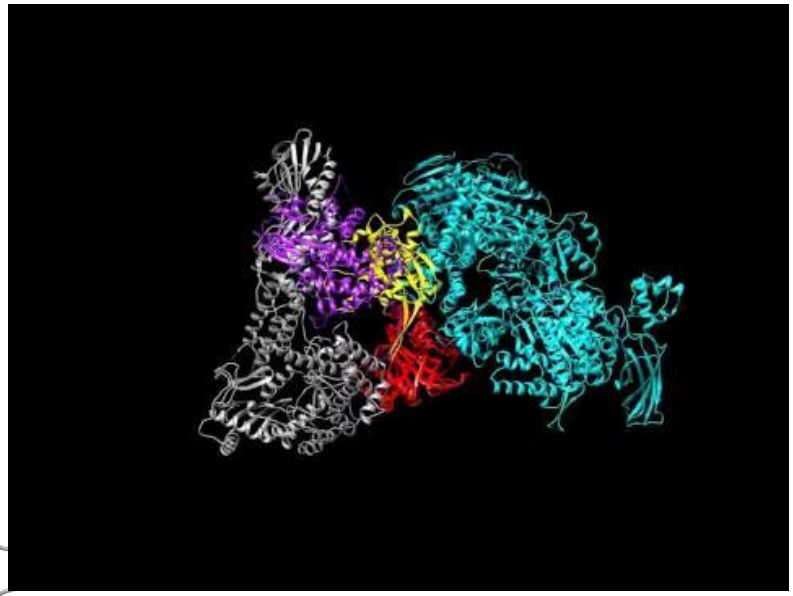
Sequence alignment based on PFAM domains

Transition Aar2-Brr2



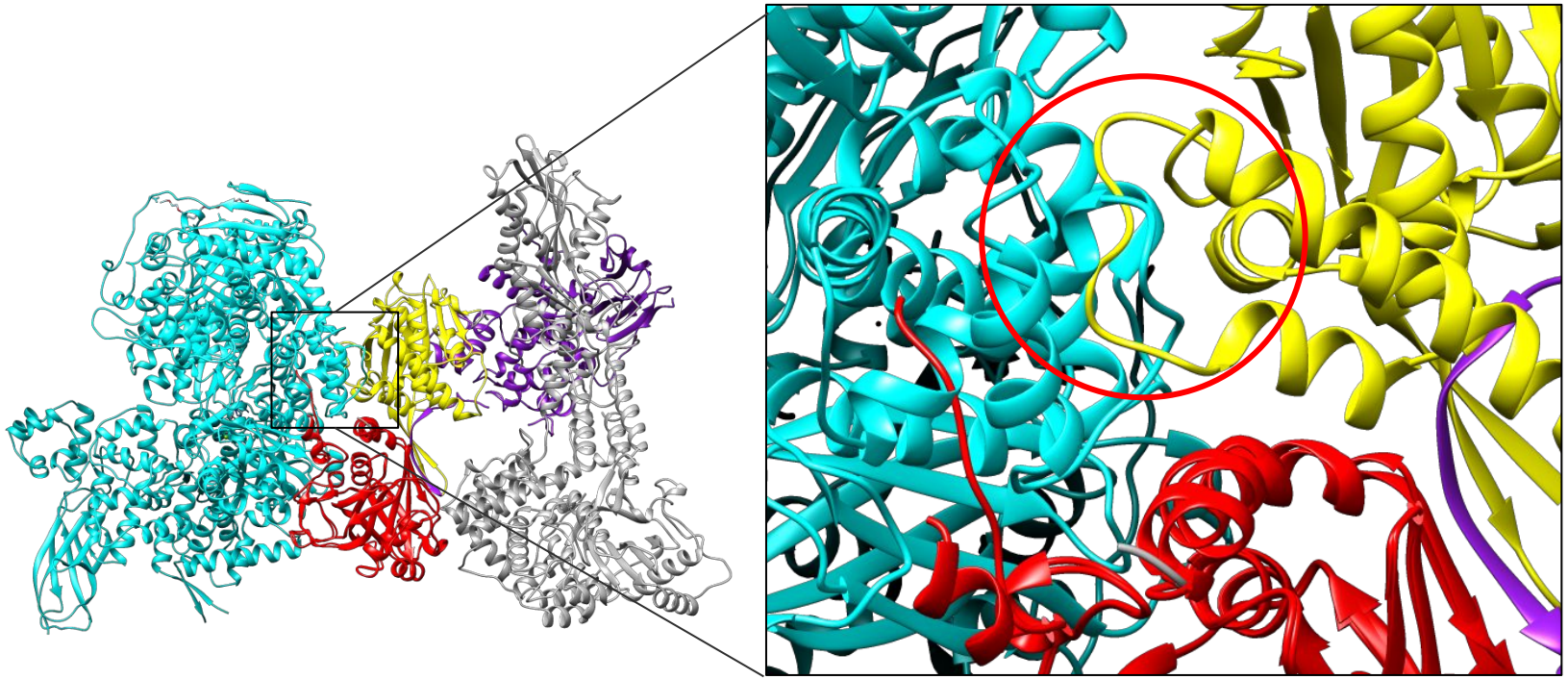
Transition Aar2-Brr2





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

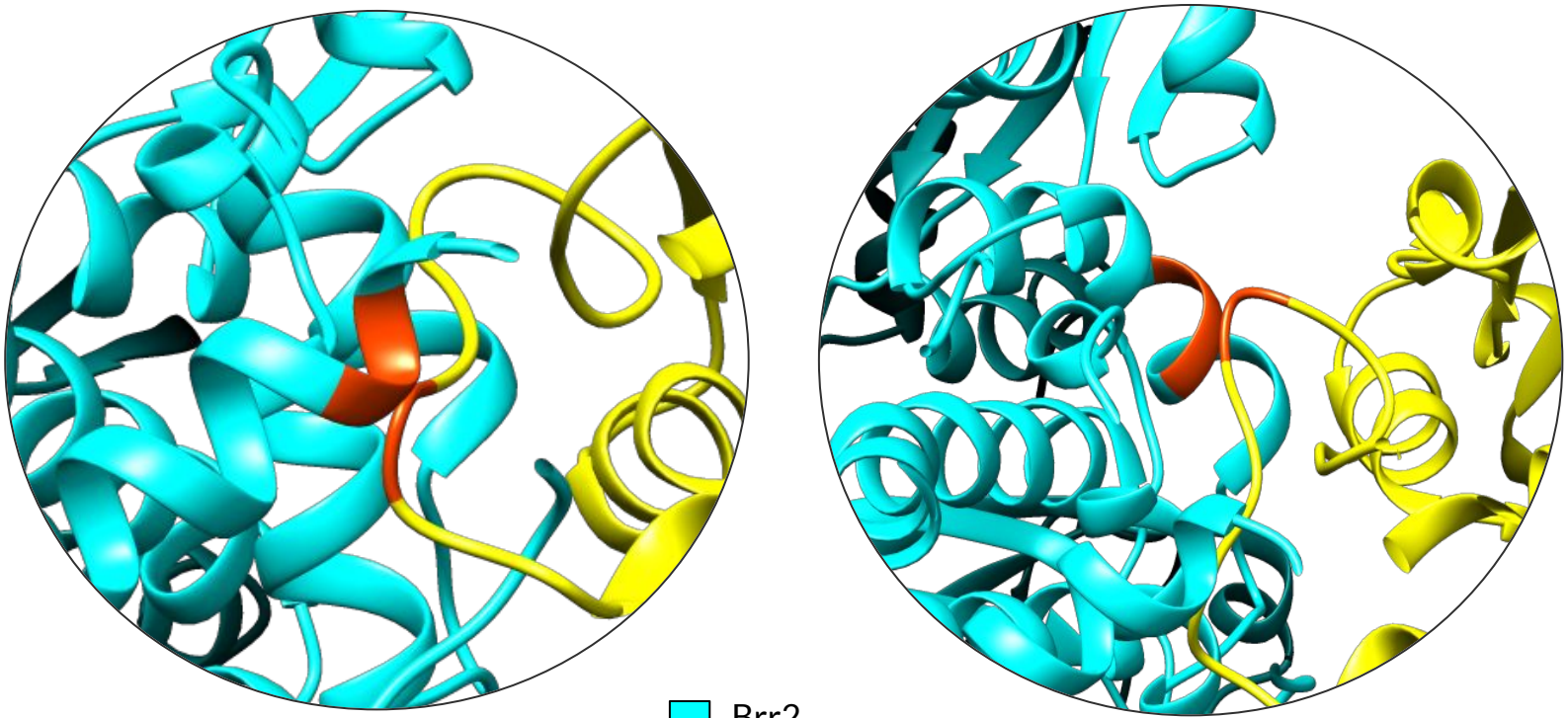
Transition Aar2-Brr2



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

Aar2
Jab1/MPN
Brr2
RnaseH-like

Transition Aar2-Brr2



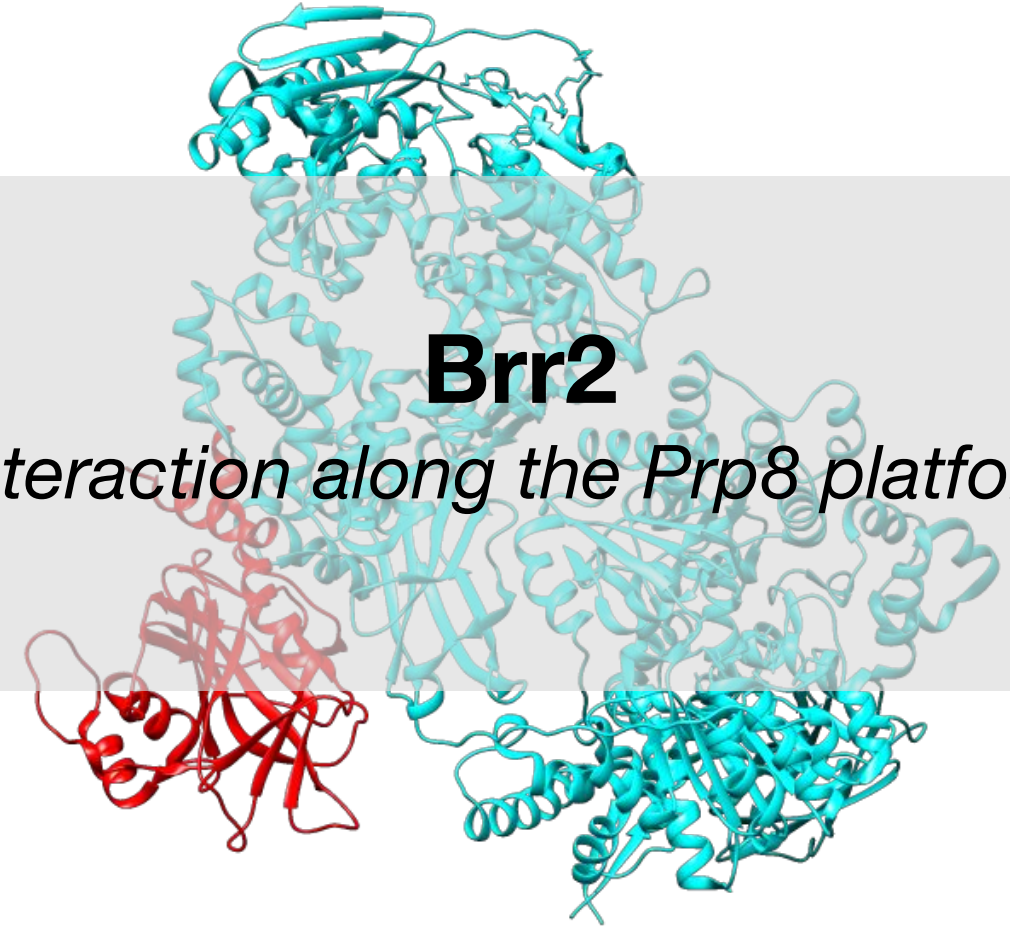
Brr2

RnaseH-like

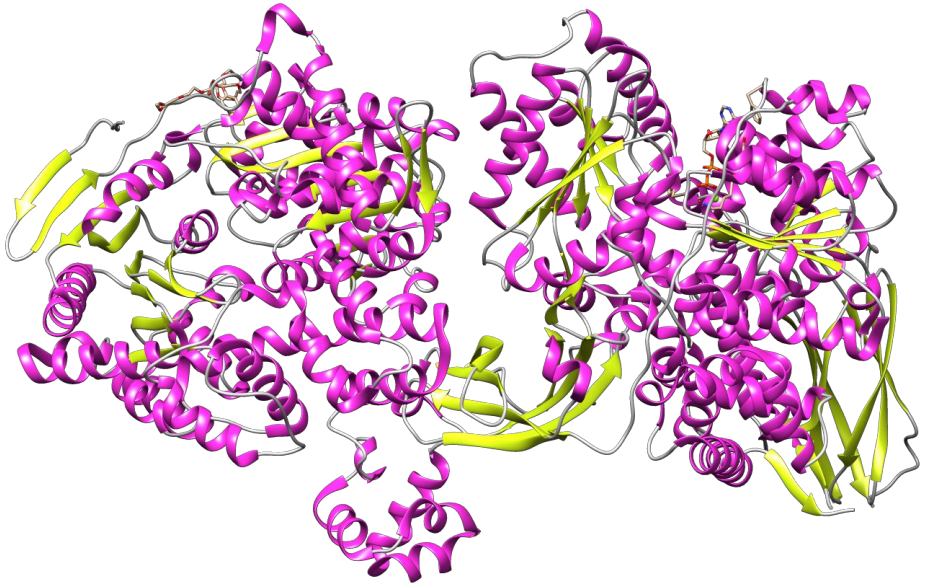
PDB code: 4I43 and 4bdg
Resolution: 2A and 2,38A

Brr2

Interaction along the Prp8 platform

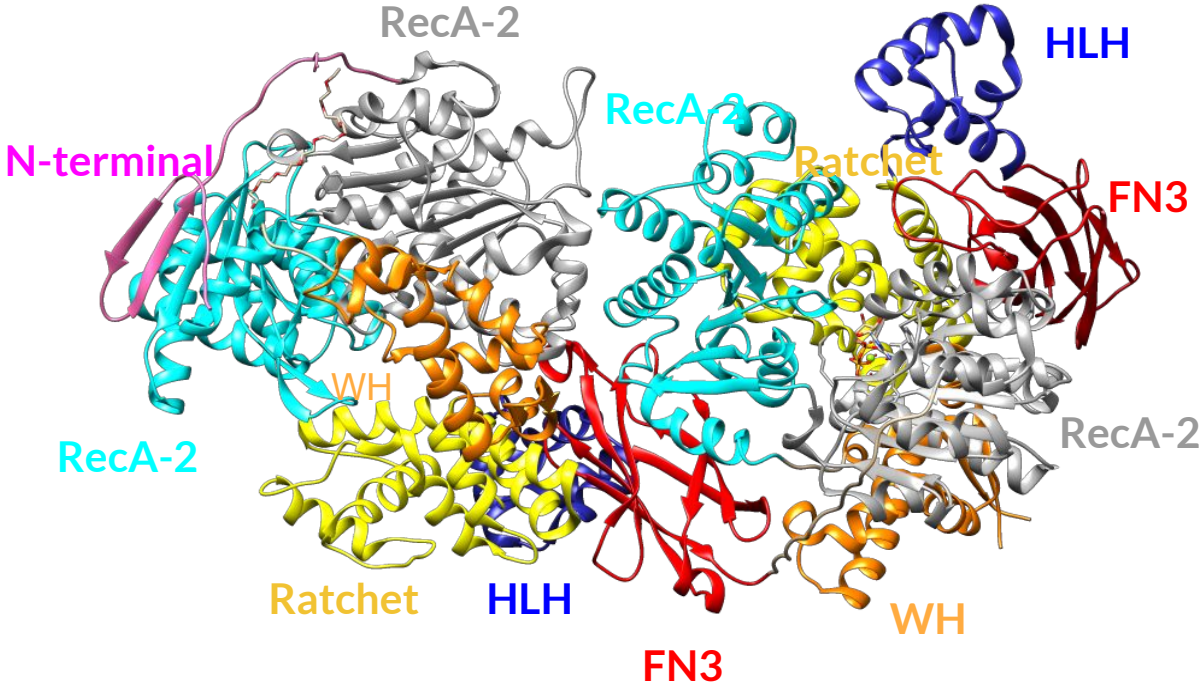


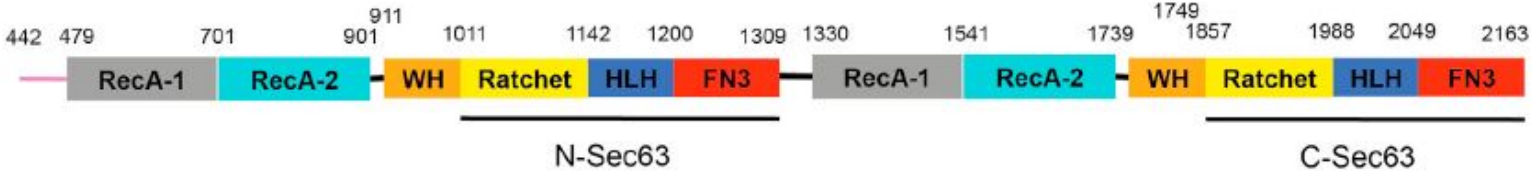
Brr2



PDB code: 4BGD
Resolution: 2,84 A

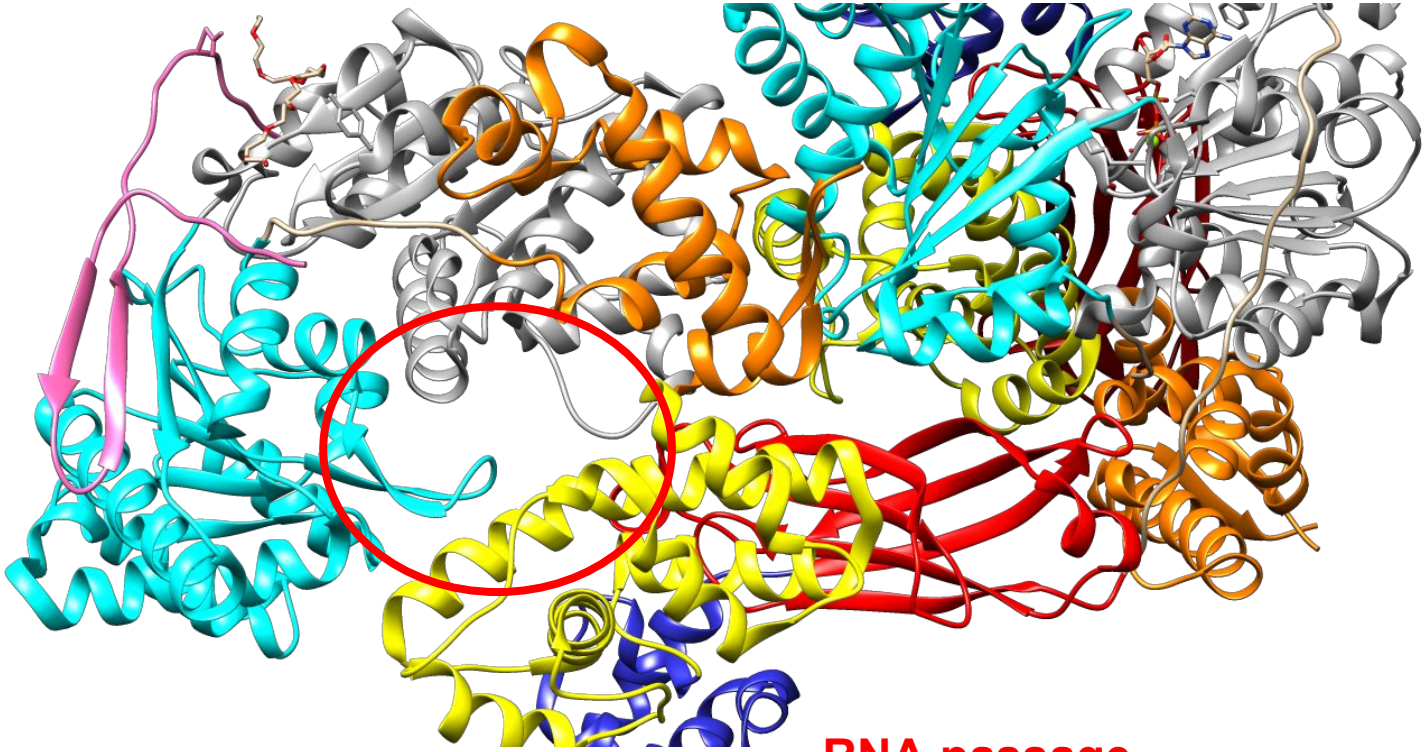
Brr2





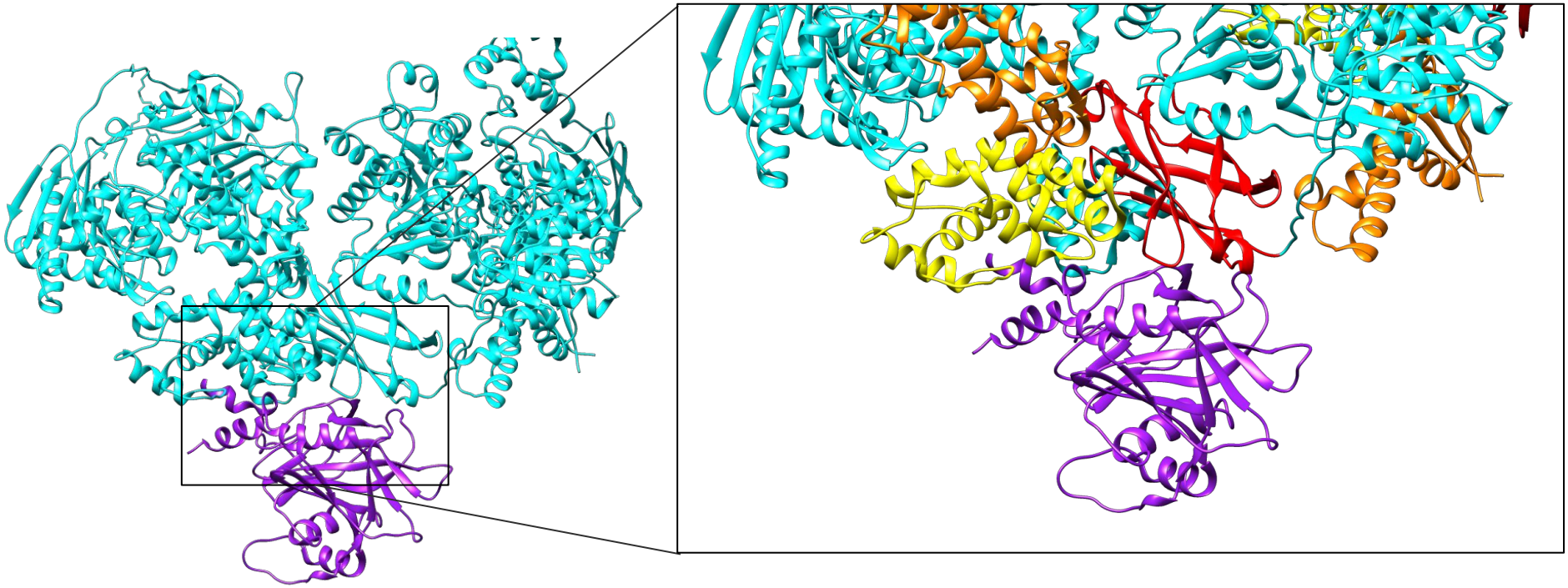
PDB code: 4bgd
Resolution: 2,84A

Brr2



PDB code: 4bgd
Resolution: 2,84A

Brr2

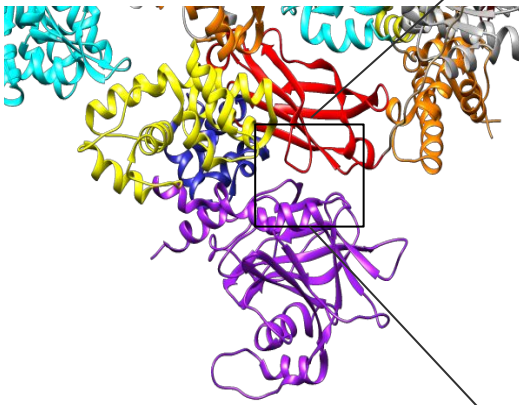


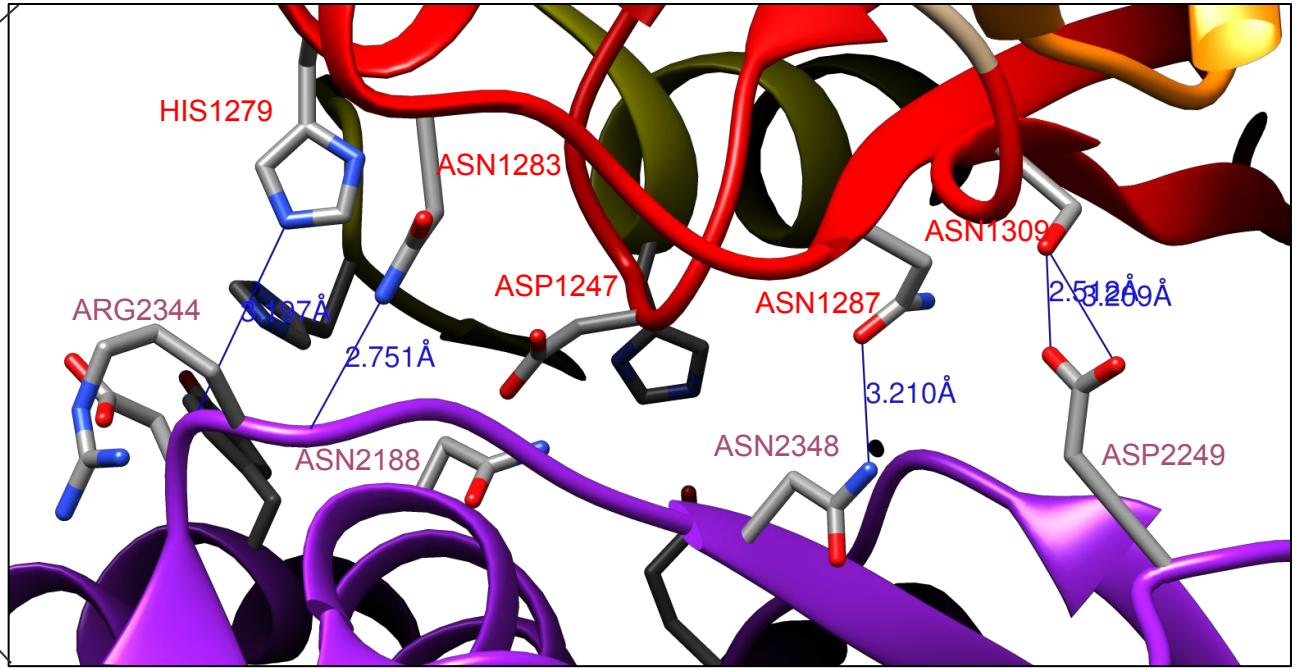
PDB code: 4bgd
Resolution: 2,84A

Brr2
FN3
Ratchet
Jab1/MPN

Brr2

Interaction Between Prp8 and the FN3 domain





PDB code: 4bgd
Resolution: 2,84A

FN3

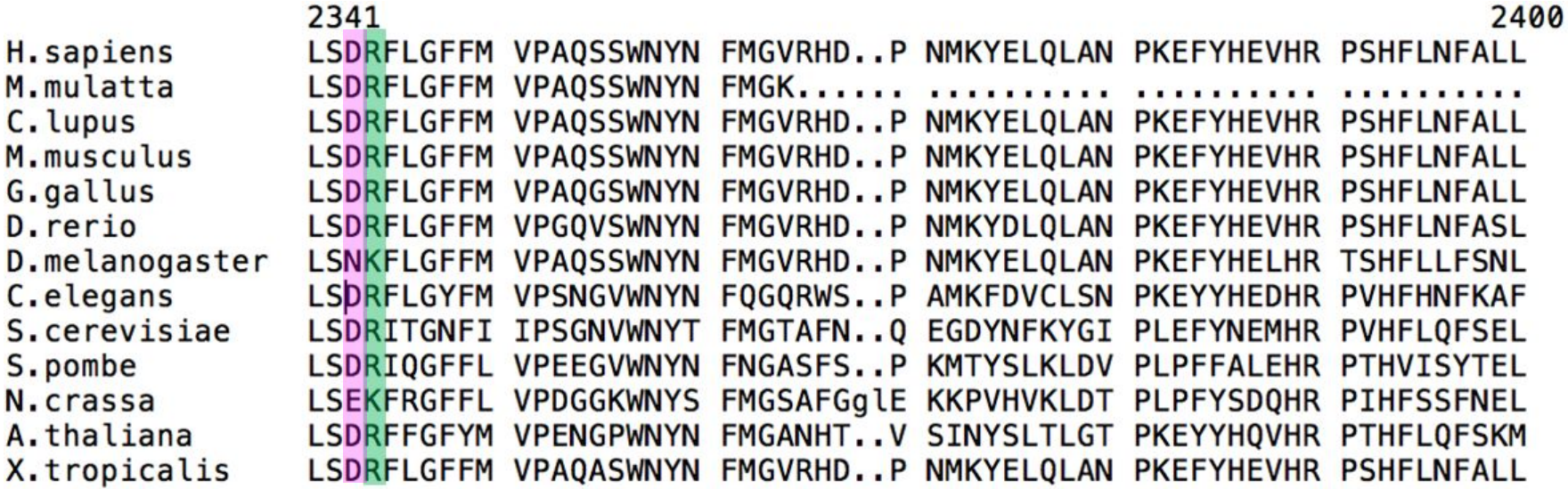
Jab1/MPN

Brr2

Asp 2343

Arg 2344

Jab1/MPN region



Negatively charged

Non-polar

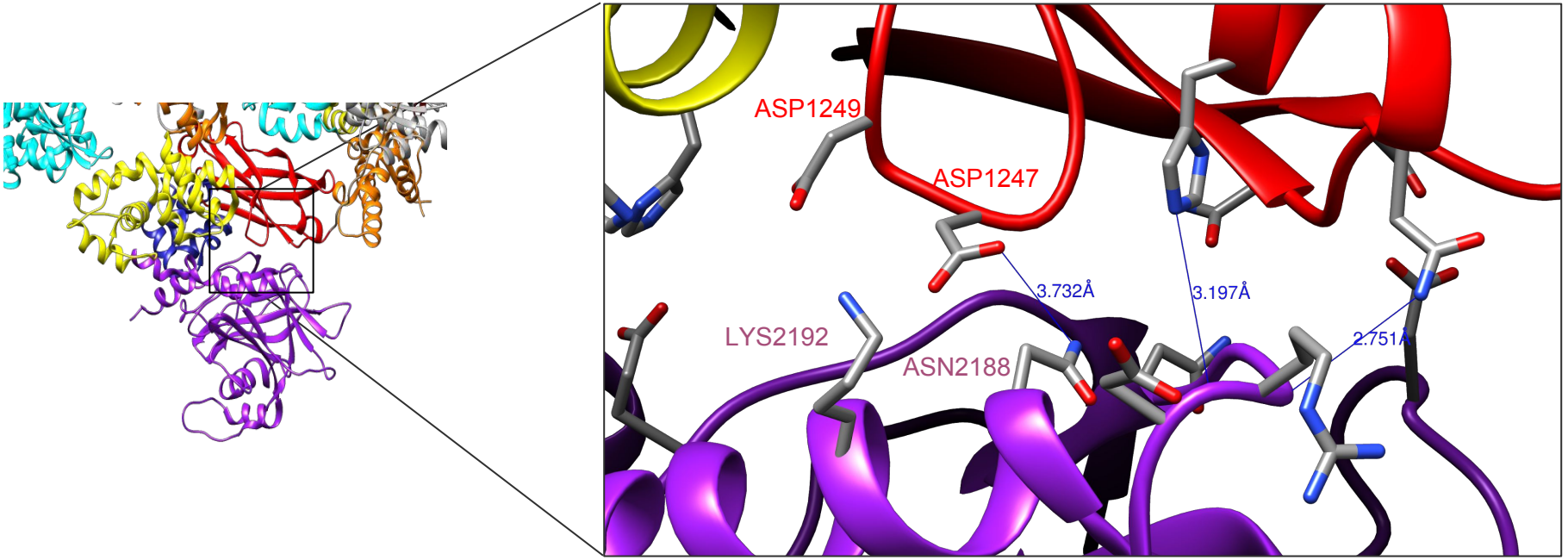
Positively charged

Polar

Sequence alignment based on PFAM domains

Brr2

Interaction Between Prp8 and the FN3 domain



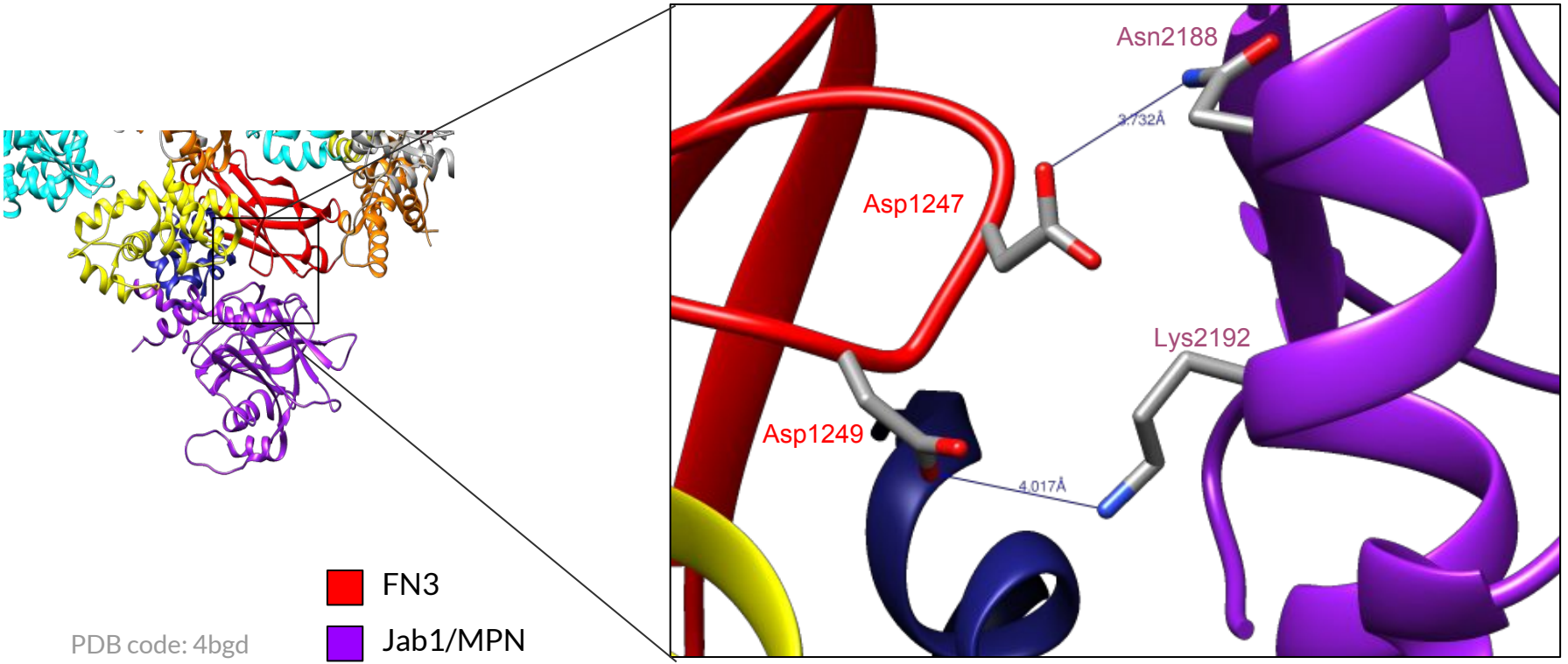
FN3

Jab1/MPN

PDB code: 4bgd
Resolution: 2,84A

Brr2

Interaction Between Prp8 and the FN3 domain



PDB code: 4bgd
Resolution: 2,84A

Brr2

Jab1/MPN region

Asn 2188

Lys 2192

```
              2161                                                           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     ITATTSNYET 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 VDSDLDD.VT YVMPKNILKR
S.cerevisiae  VVVASADYES QTFSSKNEWR KSAIANTLLY LRLKNIYVSA DDFVEEQ.NV YVLPKNLLKK
```

Negatively charged

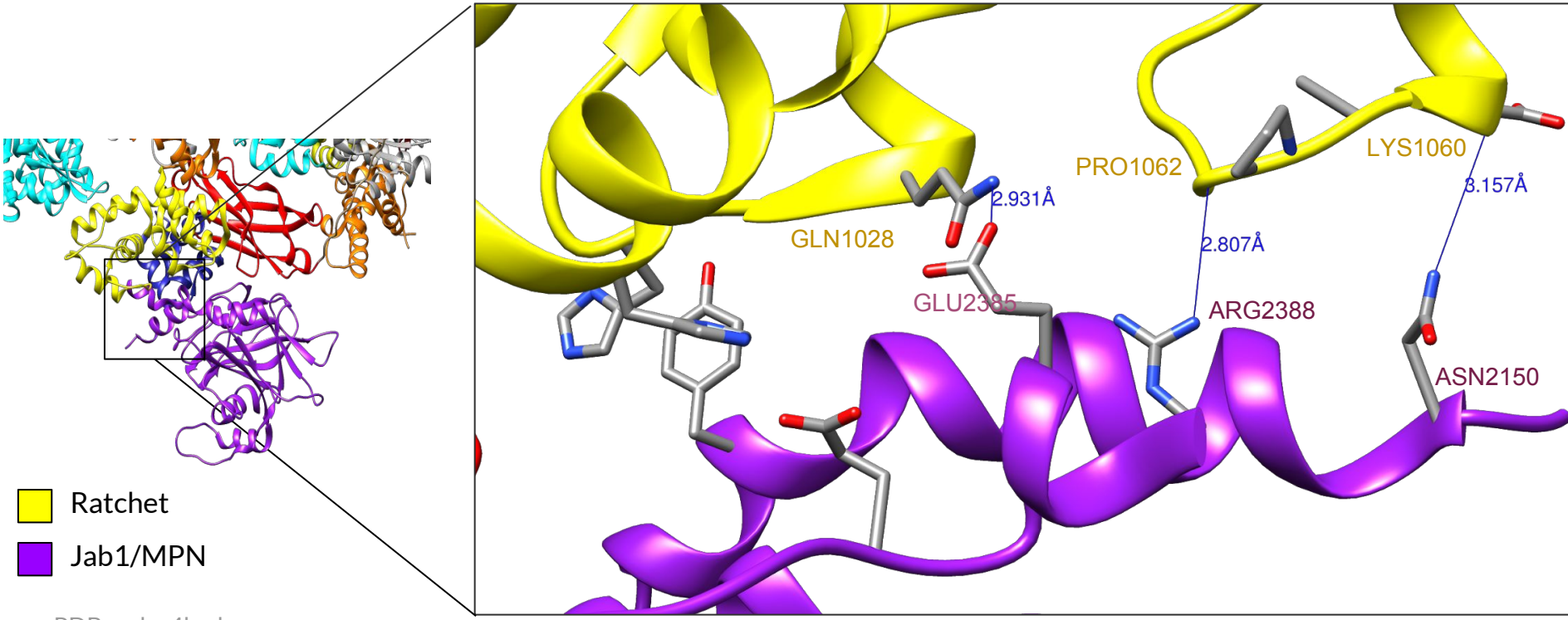
Non-polar

Positively charged

Polar

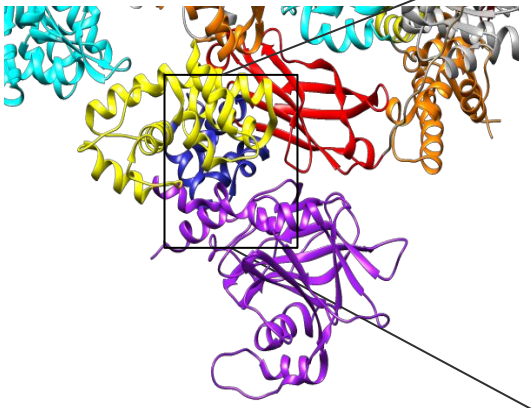
Clustal Sequence alignment

Brr2



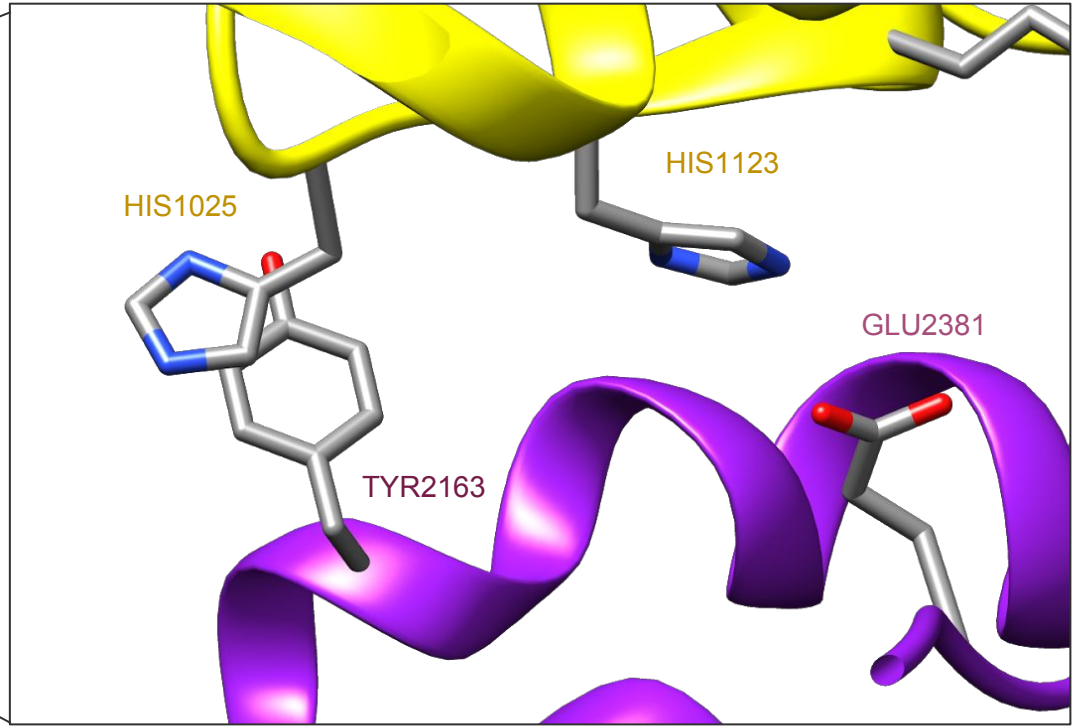
PDB code: 4bgd
Resolution: 2,84A

Brr2



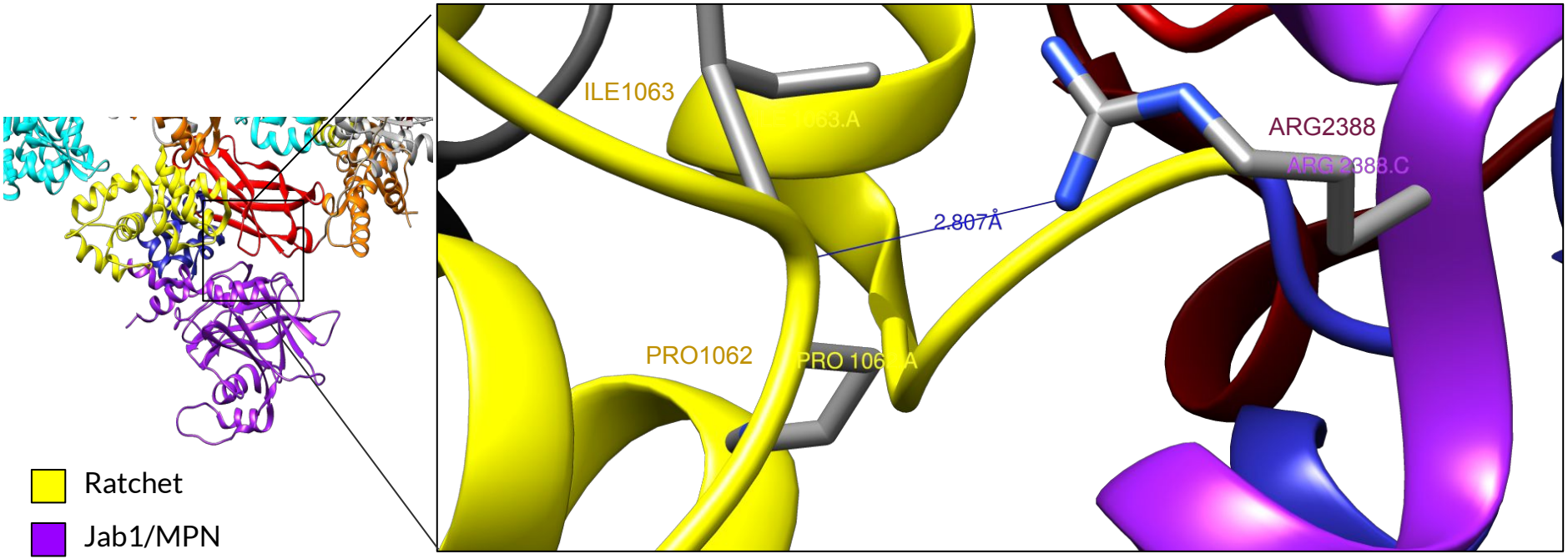
Ratchet

Jab1/MPN



PDB code: 4bgd
Resolution: 2,84A

Brr2



PDB code: 4bgd
Resolution: 2,84A

Brr2

Jab1/MPN region

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 PSHFLNFASL
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        LSEKFRGFFL VPDGGKWNYS FMGSAFGglE KKPVHVKLDT PLPFYSDQHR PIHFSSFNEL
A.thaliana      LSDRFFGFYM VPENGPWNYN FMGANHT..V SINYSLTLGT PKEYYHQVHR PTHFLQFSKM
X.tropicalis    LSDRFLGFFM VPAQASWNYN FMGVRHD..P NMKYELQLAN PKEFYHEVHR PSHFLNFALL
```

Negatively charged

Non-polar

Positively charged

Polar

*RP13

Sequence alignment based on PFAM domains

Brr2

Asn 2150

Jab1/MPN region

```
                2161                                                    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       ITATTSNYET 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 VDSDLDD.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 c**atalytic 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 A 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?

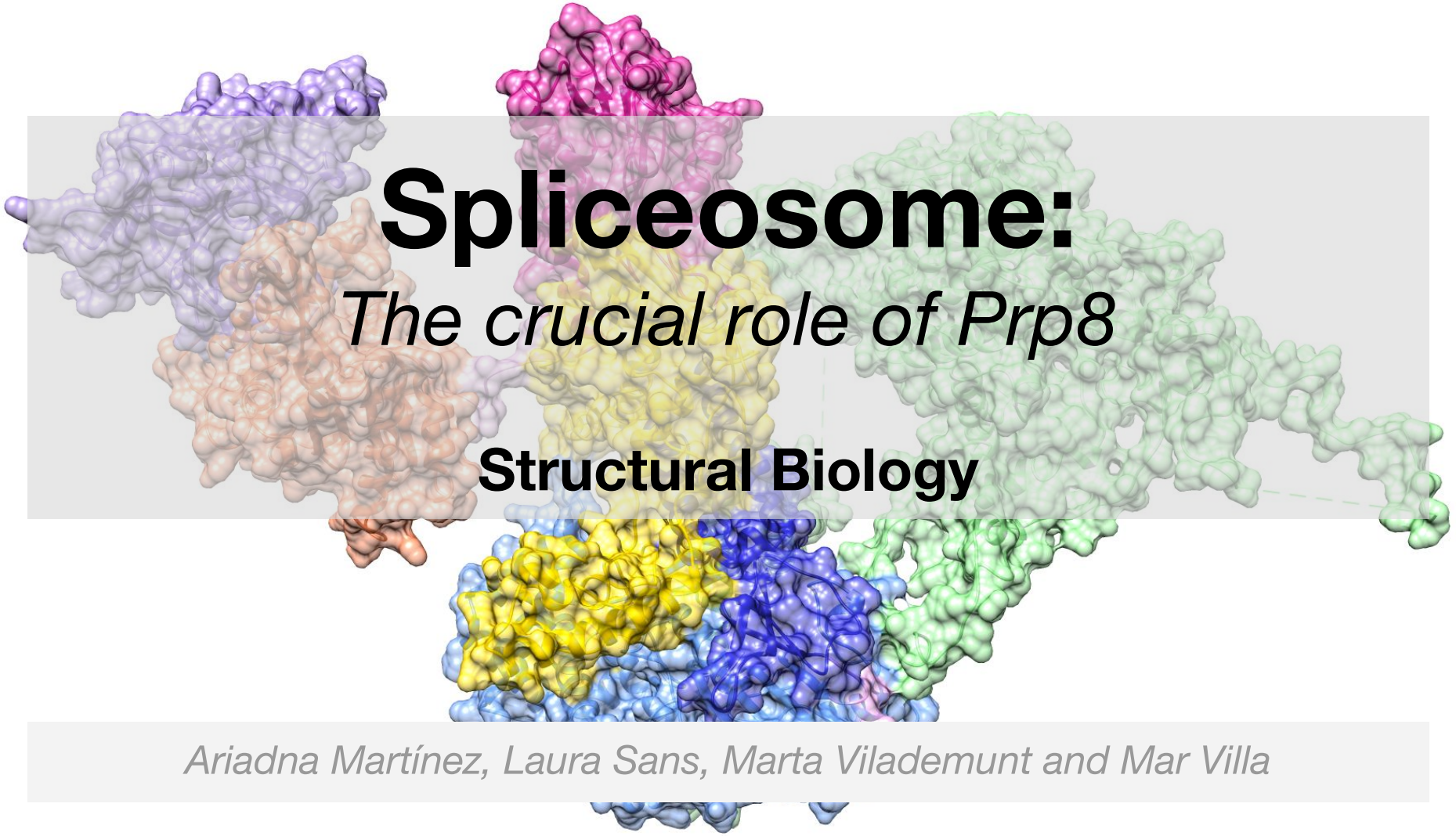
1. N-terminal domain
2. Jab1/MPN
3. RNAse H like
4. None of them

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



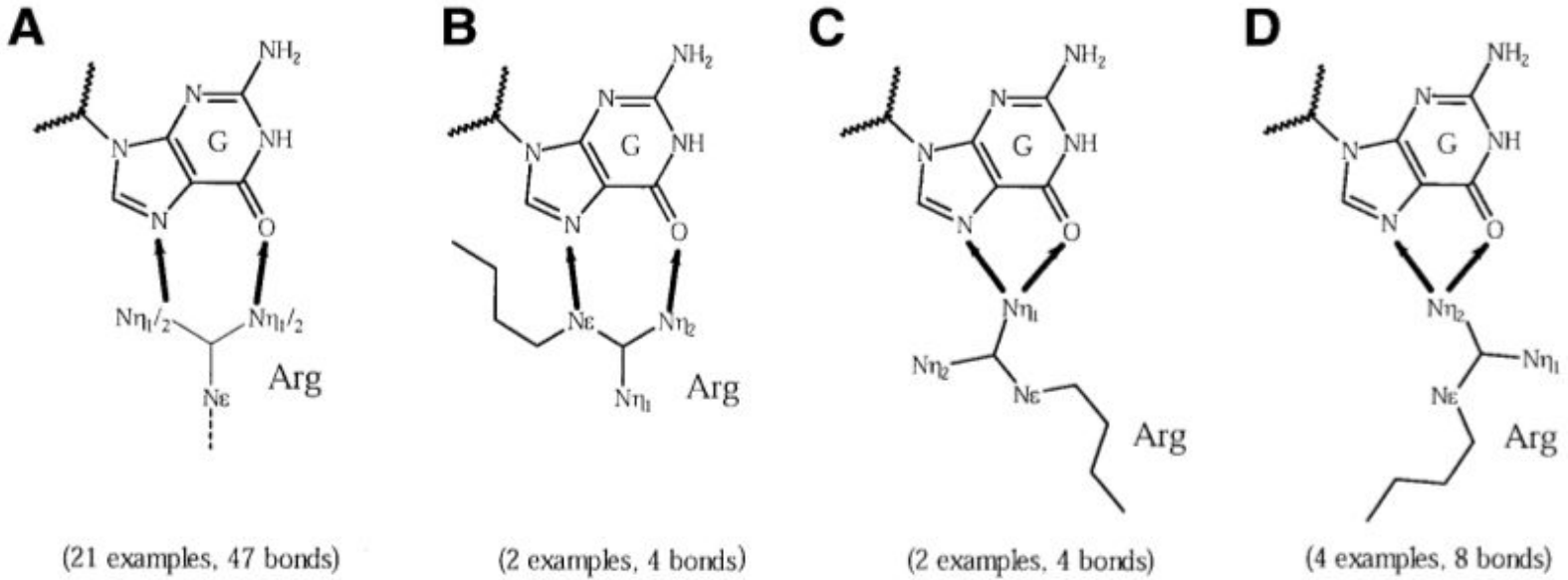
Spliceosome:

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
```



(21 examples, 47 bonds)

(2 examples, 4 bonds)

(2 examples, 4 bonds)

(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
    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)

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
```