



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:56 AM GMT

PDB ID : 2C37  
Title : RNASE PH CORE OF THE ARCHAEAL EXOSOME IN COMPLEX WITH  
U8 RNA  
Authors : Lorentzen, E.; Conti, E.  
Deposited on : 2005-10-04  
Resolution : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

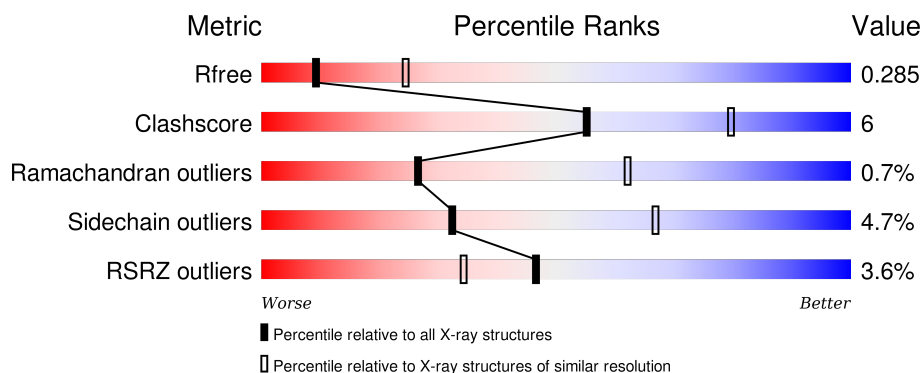
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	275	<div> <div>8%</div> <div> <div></div> <div>80%</div> <div>14%</div> <div>• 5%</div> </div> </div>
1	C	275	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>12%</div> <div>• 5%</div> </div> </div>
1	E	275	<div> <div>4%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>• 5%</div> </div> </div>
1	G	275	<div> <div>7%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>• 5%</div> </div> </div>
1	I	275	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>• 5%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	275	
1	M	275	
1	O	275	
1	Q	275	
1	S	275	
1	U	275	
1	W	275	
2	B	248	
2	D	248	
2	F	248	
2	H	248	
2	J	248	
2	L	248	
2	N	248	
2	P	248	
2	R	248	
2	T	248	
2	V	248	
2	X	248	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CL	L	301	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 46570 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called PROBABLE EXOSOME COMPLEX EXONUCLEASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			1963	1249	323	386	5			
1	C	260	Total	C	N	O	S	0	0	0
			1972	1255	325	387	5			
1	E	260	Total	C	N	O	S	0	0	0
			1968	1253	324	386	5			
1	G	260	Total	C	N	O	S	0	0	0
			1964	1251	324	384	5			
1	I	260	Total	C	N	O	S	0	0	0
			1961	1250	323	383	5			
1	K	260	Total	C	N	O	S	0	0	0
			1964	1251	324	384	5			
1	M	260	Total	C	N	O	S	0	0	0
			1975	1256	325	389	5			
1	O	260	Total	C	N	O	S	0	0	0
			1968	1253	325	385	5			
1	Q	260	Total	C	N	O	S	0	0	0
			1968	1253	325	385	5			
1	S	260	Total	C	N	O	S	0	0	0
			1967	1253	324	385	5			
1	U	260	Total	C	N	O	S	0	0	0
			1968	1253	325	385	5			
1	W	260	Total	C	N	O	S	0	0	0
			1975	1256	324	390	5			

- Molecule 2 is a protein called PROBABLE EXOSOME COMPLEX EXONUCLEASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	241	Total	C	N	O	S	0	0	0
			1859	1172	321	356	10			
2	D	247	Total	C	N	O	S	0	0	0
			1901	1199	328	363	11			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	241	Total	C	N	O	S	0	0	0
			1851	1168	321	352	10			
2	H	241	Total	C	N	O	S	0	0	0
			1859	1172	321	356	10			
2	J	241	Total	C	N	O	S	0	0	0
			1859	1172	321	356	10			
2	L	247	Total	C	N	O	S	0	0	0
			1905	1201	328	365	11			
2	N	241	Total	C	N	O	S	0	0	0
			1859	1172	321	356	10			
2	P	248	Total	C	N	O	S	0	0	0
			1903	1201	326	364	12			
2	R	245	Total	C	N	O	S	0	0	0
			1885	1189	326	359	11			
2	T	247	Total	C	N	O	S	0	0	0
			1901	1199	328	363	11			
2	V	244	Total	C	N	O	S	0	0	0
			1879	1186	324	359	10			
2	X	241	Total	C	N	O	S	0	0	0
			1859	1172	321	356	10			

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

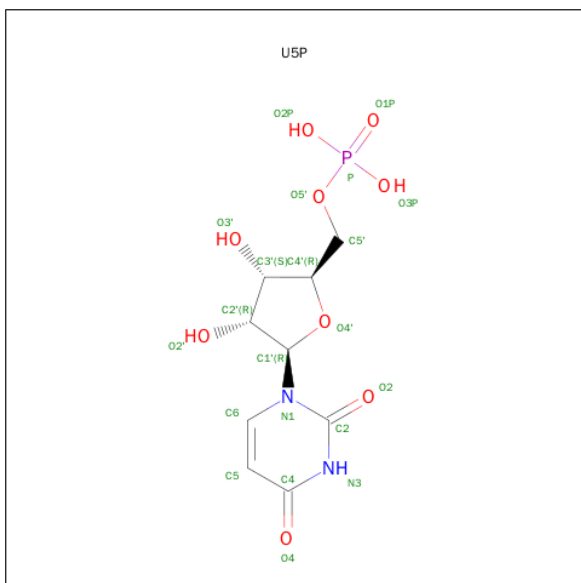
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total	Cl	0	0
			1	1		
3	J	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		
3	H	1	Total	Cl	0	0
			1	1		
3	B	1	Total	Cl	0	0
			1	1		
3	V	1	Total	Cl	0	0
			1	1		
3	T	1	Total	Cl	0	0
			1	1		
3	N	1	Total	Cl	0	0
			1	1		
3	X	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	R	1	Total	Cl	0	0
			1	1		
3	L	1	Total	Cl	0	0
			1	1		
3	F	1	Total	Cl	0	0
			1	1		

- Molecule 4 is URIDINE-5'-MONOPHOSPHATE (three-letter code: U5P) (formula:  $C_9H_{13}N_2O_9P$ ).



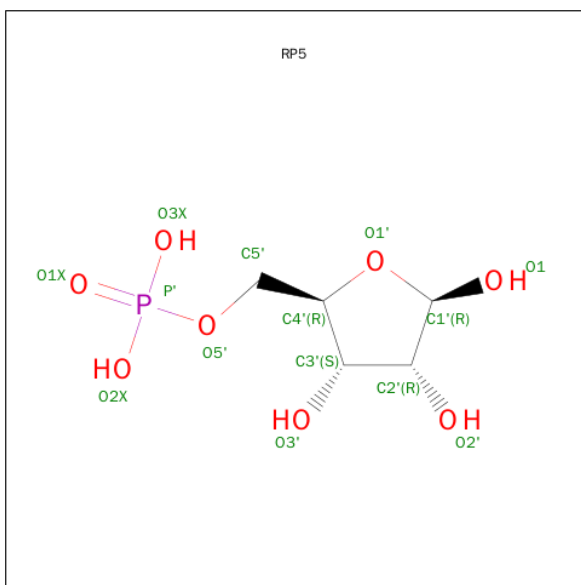
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	0	0
			21	9	2	9	1		
4	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
4	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
4	B	1	Total	O	P			0	0
			4	3	1				
4	F	1	Total	C	N	O	P	0	0
			21	9	2	9	1		
4	F	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
4	F	1	Total	C	O	P		0	0
			12	5	6	1			
4	F	1	Total	C	O	P		0	0
			12	5	6	1			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	1	Total C O P 13 5 7 1	0	0
4	H	1	Total C N O P 20 9 2 8 1	0	0
4	H	1	Total C N O P 20 9 2 8 1	0	0
4	J	1	Total C N O P 21 9 2 9 1	0	0
4	J	1	Total C N O P 20 9 2 8 1	0	0
4	J	1	Total O P 4 3 1	0	0
4	N	1	Total C N O P 21 9 2 9 1	0	0
4	N	1	Total C O P 12 5 6 1	0	0
4	N	1	Total C O P 12 5 6 1	0	0
4	V	1	Total C N O P 21 9 2 9 1	0	0
4	V	1	Total C N O P 20 9 2 8 1	0	0
4	V	1	Total C N O P 20 9 2 8 1	0	0
4	V	1	Total C N O P 20 9 2 8 1	0	0
4	X	1	Total C N O P 21 9 2 9 1	0	0
4	X	1	Total C N O P 20 9 2 8 1	0	0
4	X	1	Total C O P 12 5 6 1	0	0
4	X	1	Total O P 4 3 1	0	0

- Molecule 5 is 5-O-PHOSPHONO-BETA-D-RIBOFURANOSE (three-letter code: RP5) (formula: C<sub>5</sub>H<sub>11</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	N	1	Total	C	O	P	0	0
			13	5	7	1		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

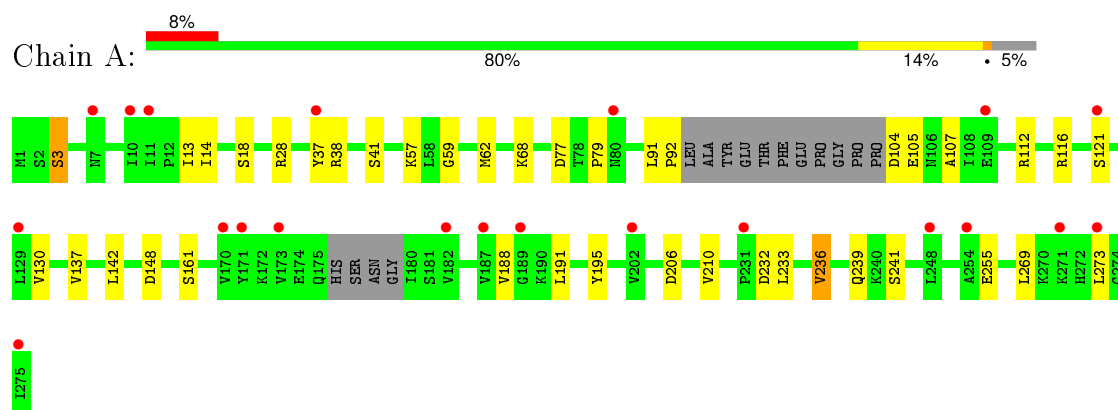
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	V	1	Total	Na	0	0
			1	1		



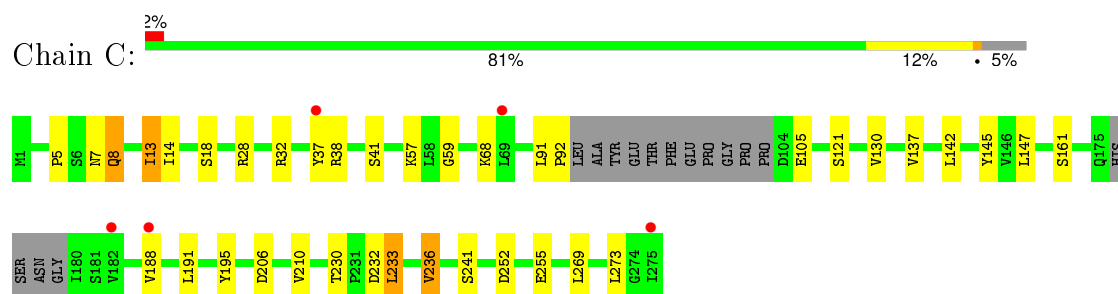
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $\text{RSRZ} > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

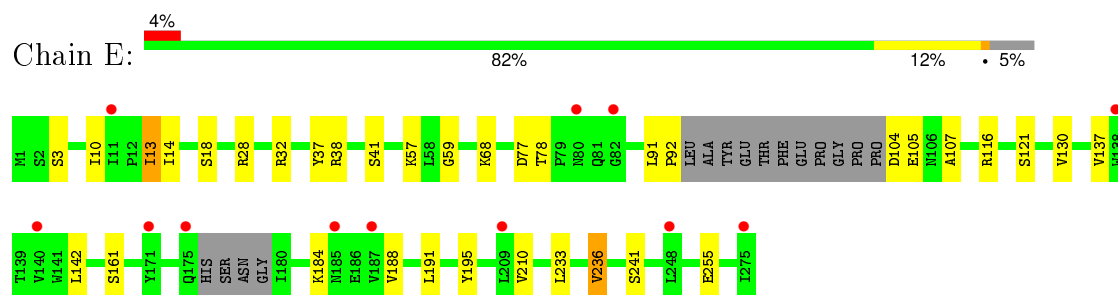
#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



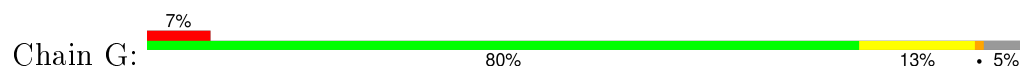
#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2

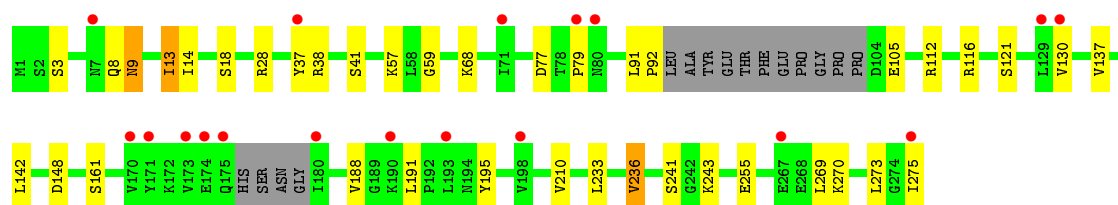


#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2

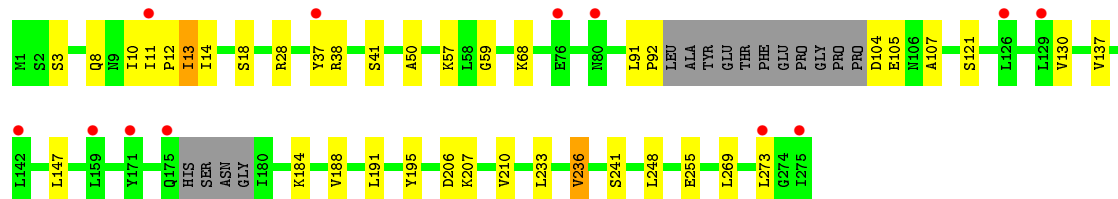
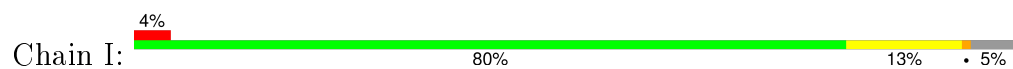


#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2

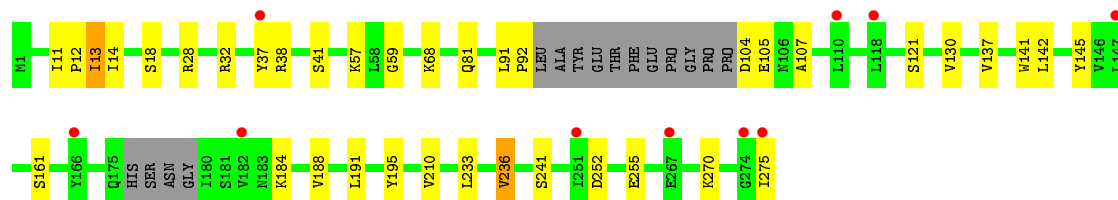
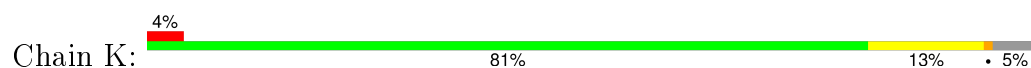




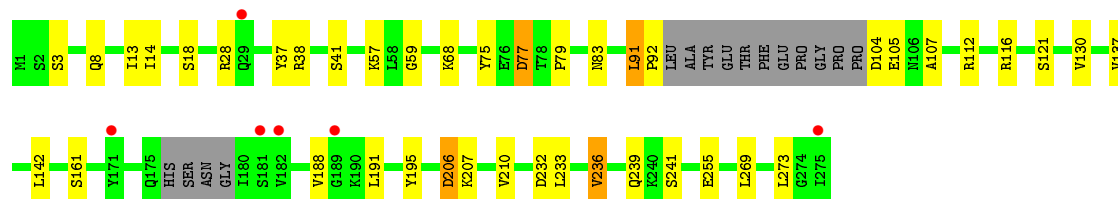
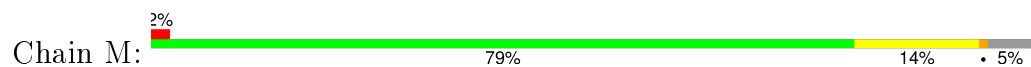
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



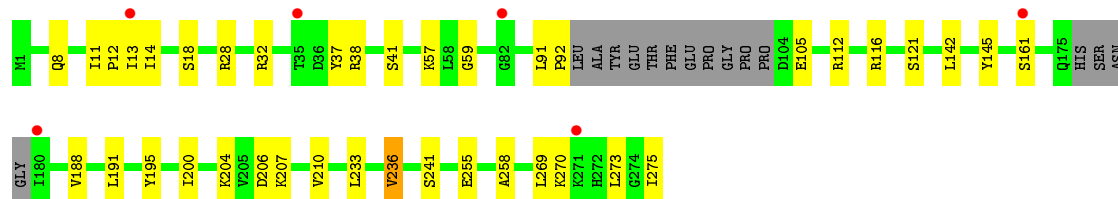
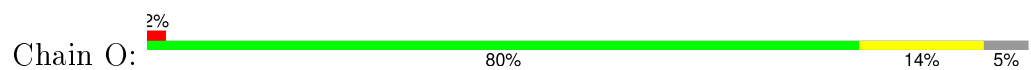
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



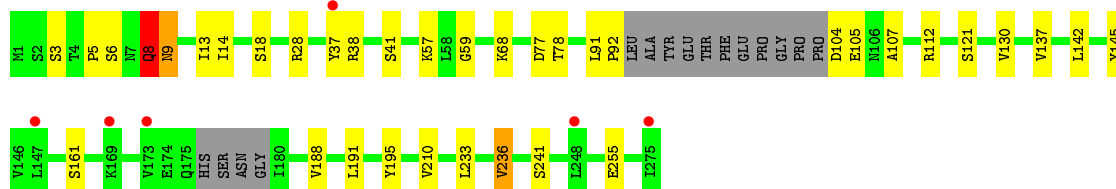
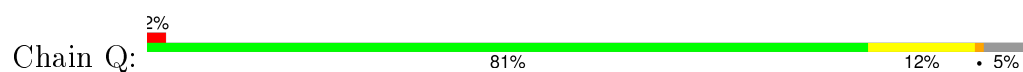
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



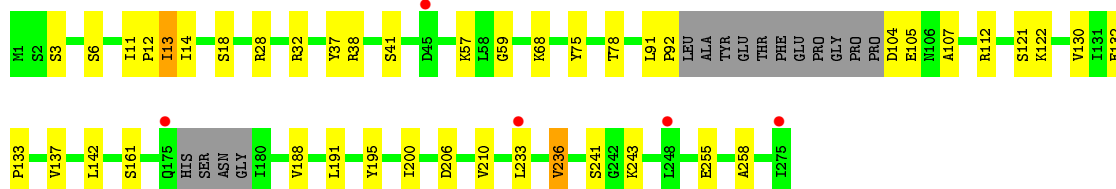
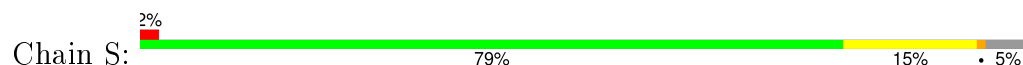
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



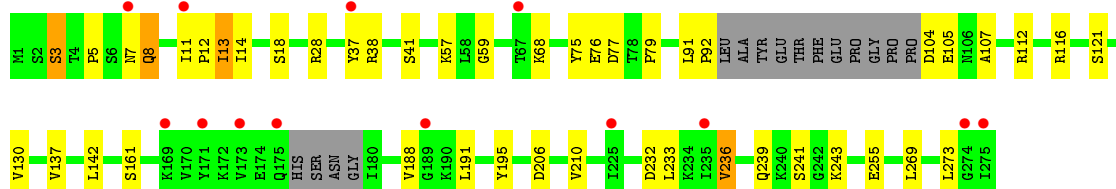
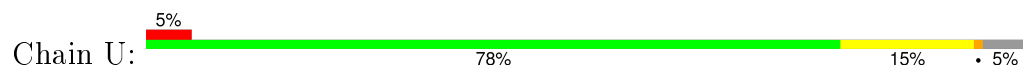
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



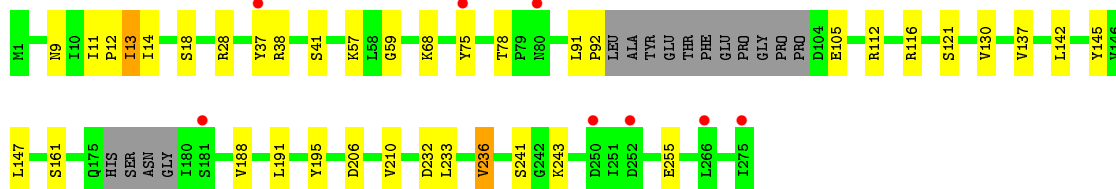
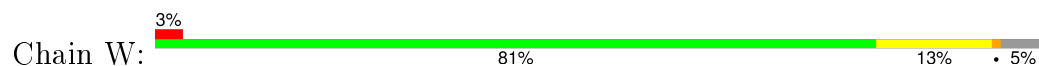
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



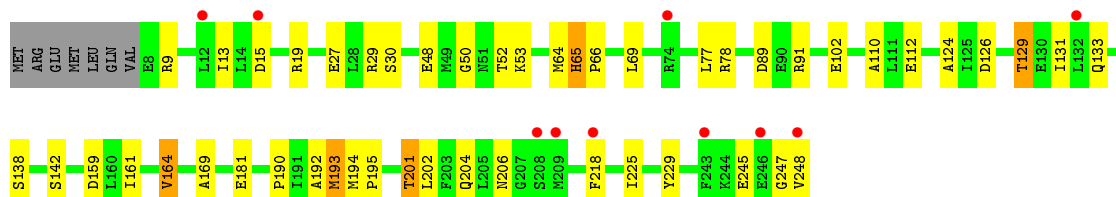
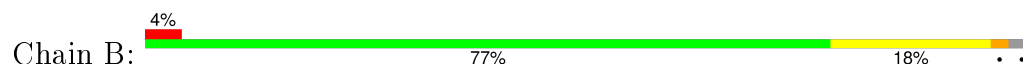
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



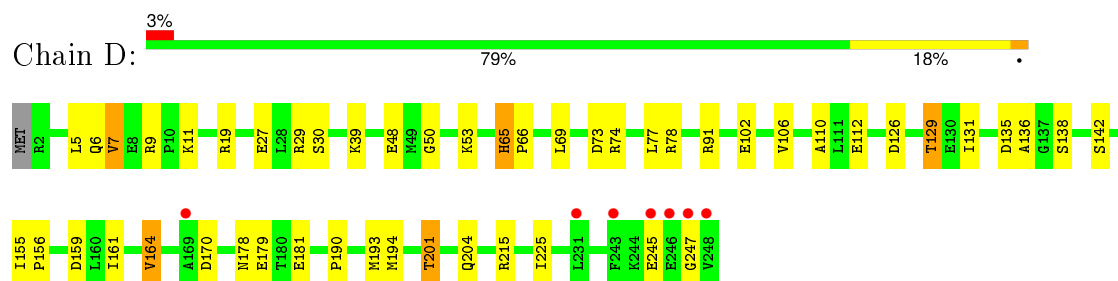
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



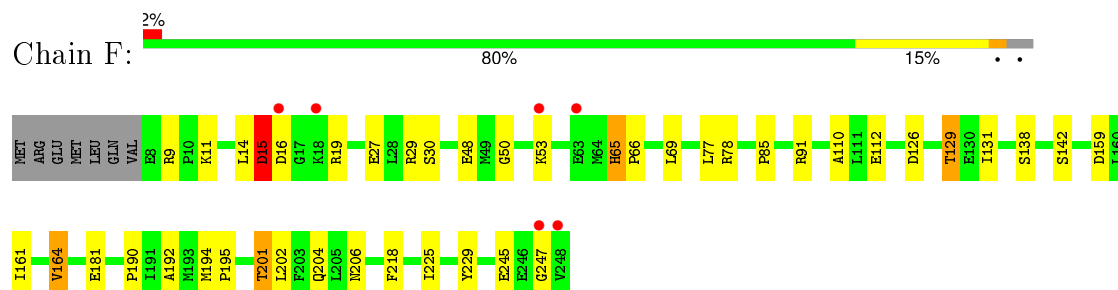
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



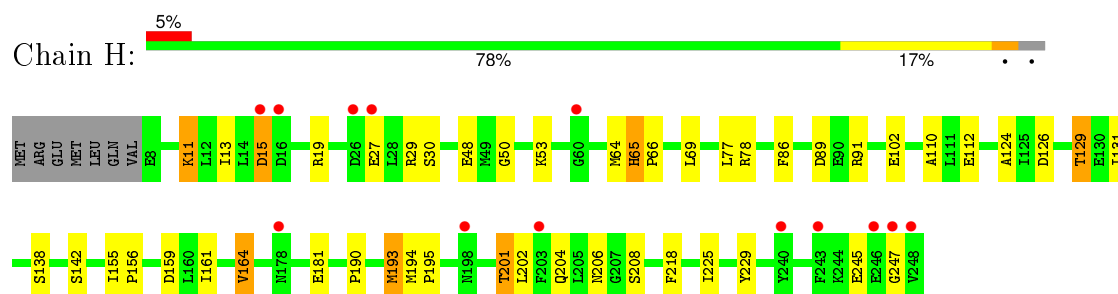
- Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



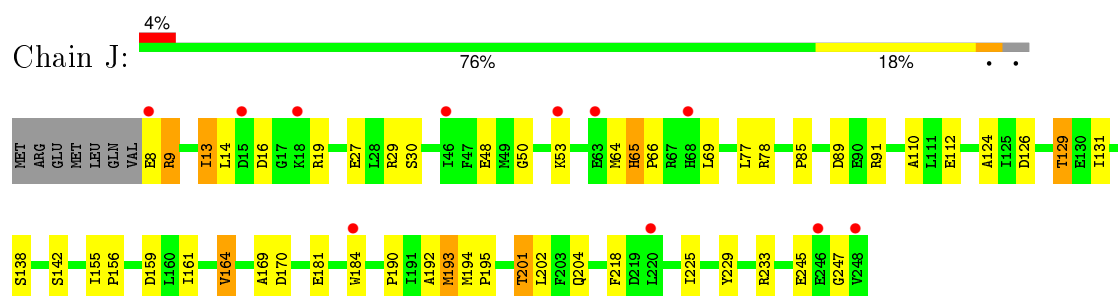
- Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



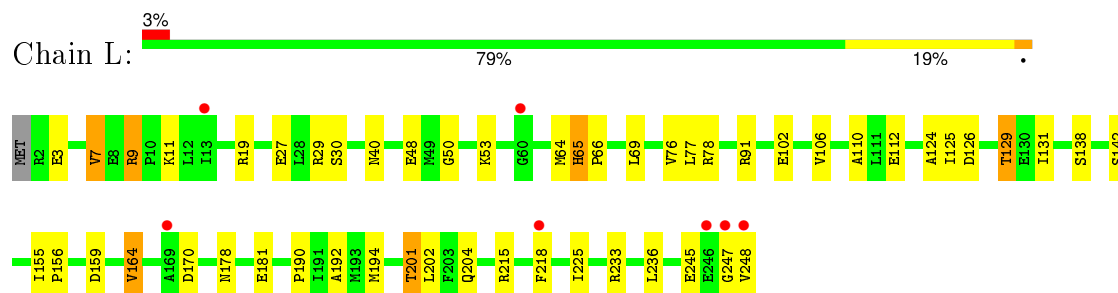
- Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



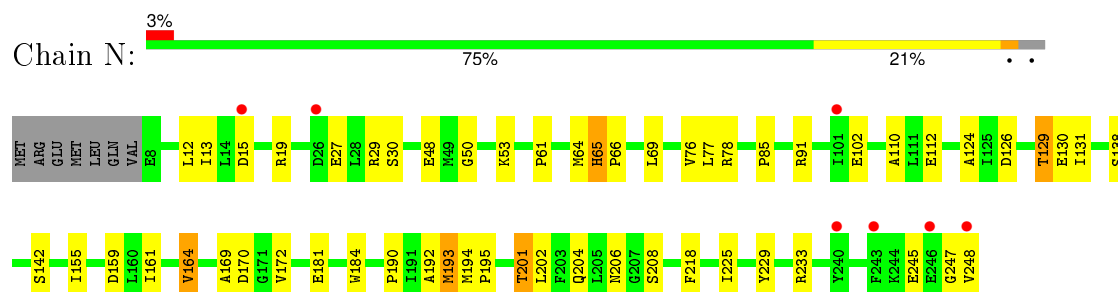
- Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



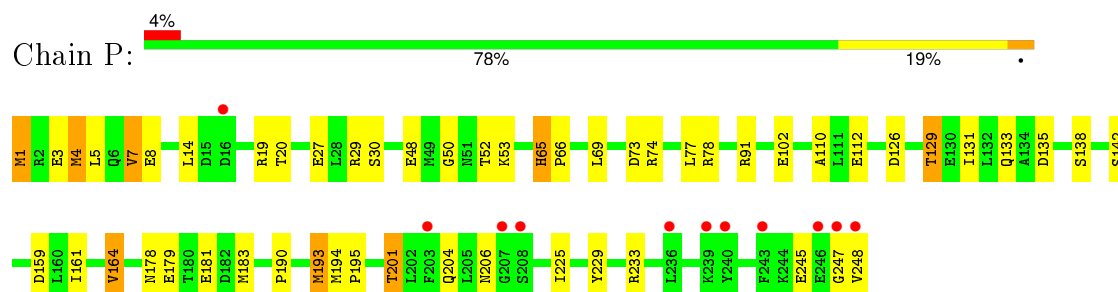
- Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



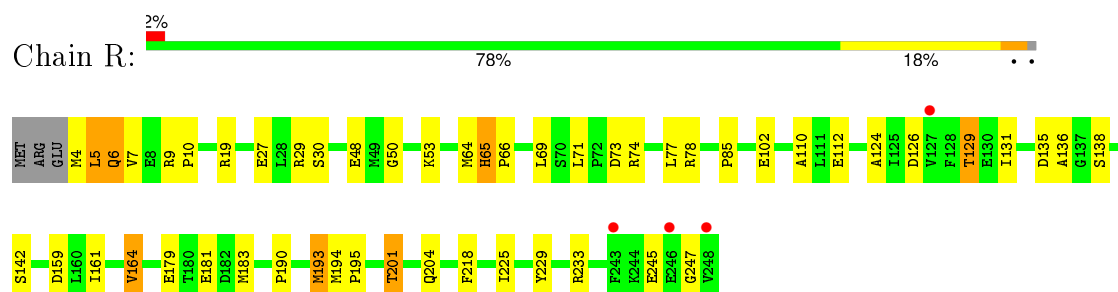
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



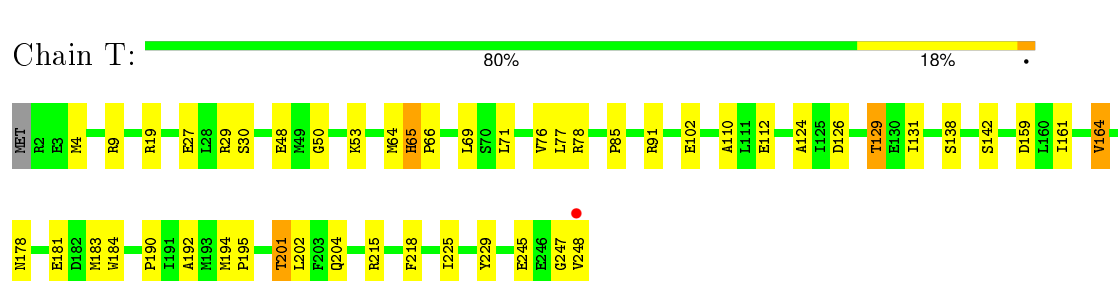
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



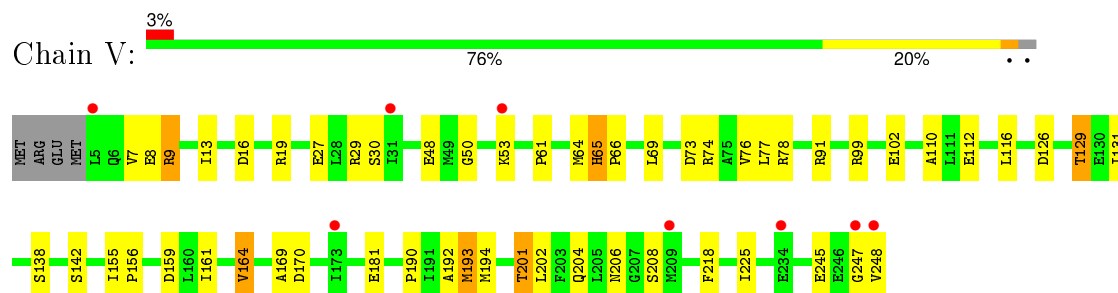
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



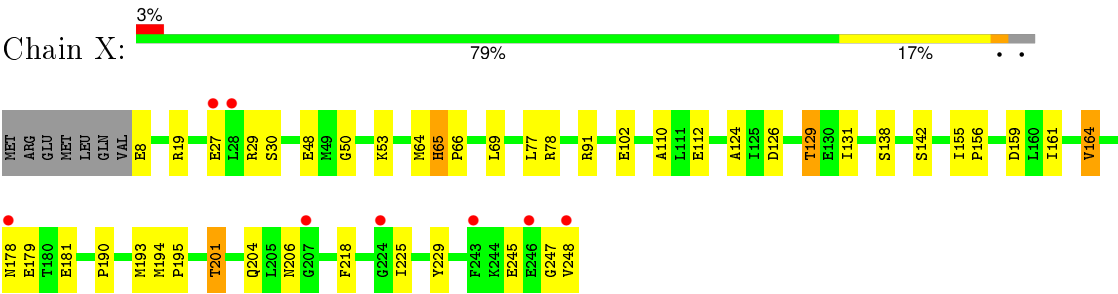
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



● Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	207.02Å 212.94Å 433.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	93.25 – 2.80 87.47 – 2.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (93.25-2.80) 93.3 (87.47-2.80)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 2.82Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.266 , 0.282 0.272 , 0.285	Depositor DCC
$R_{free}$ test set	6530 reflections (3.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	57.0	Xtriage
Anisotropy	0.195	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 17.6	EDS
Estimated twinning fraction	0.045 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	1 of 217623 reflections (0.000%)	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	46570	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, RP5, U5P, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.45	0/1988	0.57	0/2701
1	C	0.50	0/1997	0.59	0/2711
1	E	0.47	0/1993	0.57	0/2706
1	G	0.46	0/1989	0.58	0/2701
1	I	0.45	0/1986	0.58	0/2697
1	K	0.48	0/1989	0.59	0/2701
1	M	0.46	0/2000	0.58	0/2715
1	O	0.50	0/1993	0.59	0/2706
1	Q	0.50	0/1993	0.59	0/2706
1	S	0.50	0/1992	0.59	0/2706
1	U	0.48	0/1993	0.59	0/2706
1	W	0.46	0/2000	0.57	0/2715
2	B	0.46	0/1887	0.61	0/2549
2	D	0.48	0/1929	0.64	0/2606
2	F	0.49	0/1879	0.61	0/2539
2	H	0.47	0/1887	0.62	0/2549
2	J	0.47	0/1887	0.60	0/2549
2	L	0.50	0/1933	0.64	0/2611
2	N	0.49	0/1887	0.62	0/2549
2	P	0.52	0/1931	0.64	0/2609
2	R	0.52	0/1913	0.63	0/2584
2	T	0.51	0/1929	0.64	0/2606
2	V	0.47	0/1907	0.62	0/2577
2	X	0.48	0/1887	0.63	0/2549
All	All	0.48	0/46769	0.60	0/63348

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.



## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1963	0	1998	22	0
1	C	1972	0	2017	21	0
1	E	1968	0	2011	19	0
1	G	1964	0	2007	19	0
1	I	1961	0	2003	21	0
1	K	1964	0	2007	20	1
1	M	1975	0	2019	22	0
1	O	1968	0	2013	20	0
1	Q	1968	0	2013	21	0
1	S	1967	0	2011	20	1
1	U	1968	0	2013	25	0
1	W	1975	0	2017	20	0
2	B	1859	0	1884	31	0
2	D	1901	0	1925	31	1
2	F	1851	0	1876	29	0
2	H	1859	0	1884	33	0
2	J	1859	0	1884	36	0
2	L	1905	0	1929	29	0
2	N	1859	0	1884	39	0
2	P	1903	0	1926	33	0
2	R	1885	0	1910	33	1
2	T	1901	0	1925	30	0
2	V	1879	0	1906	35	0
2	X	1859	0	1884	26	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
3	J	1	0	0	0	0
3	L	1	0	0	0	0
3	N	1	0	0	0	0
3	P	1	0	0	0	0
3	R	1	0	0	0	0
3	T	1	0	0	0	0
3	V	1	0	0	0	0
3	X	1	0	0	0	0
4	B	65	0	30	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	65	0	33	1	0
4	H	53	0	27	1	0
4	J	45	0	20	0	0
4	N	45	0	23	5	0
4	V	81	0	41	2	0
4	X	57	0	26	0	0
5	N	13	0	7	0	0
6	V	1	0	0	0	0
All	All	46570	0	47153	586	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 586 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:N:401:U5P:H6	4:N:401:U5P:H5'1	1.37	1.05
2:X:77:LEU:HD12	2:X:112:GLU:HG3	1.55	0.88
2:N:77:LEU:HD12	2:N:112:GLU:HG3	1.59	0.85
2:P:77:LEU:HD12	2:P:112:GLU:HG3	1.60	0.84
2:H:77:LEU:HD12	2:H:112:GLU:HG3	1.60	0.84

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:6:GLN:O	1:K:81:GLN:NE2[8_545]	2.02	0.18
2:R:9:ARG:O	1:S:122:LYS:NZ[5_455]	2.19	0.01

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	254/275 (92%)	244 (96%)	9 (4%)	1 (0%)	39	74
1	C	254/275 (92%)	244 (96%)	9 (4%)	1 (0%)	39	74
1	E	254/275 (92%)	243 (96%)	9 (4%)	2 (1%)	24	58
1	G	254/275 (92%)	245 (96%)	7 (3%)	2 (1%)	24	58
1	I	254/275 (92%)	244 (96%)	7 (3%)	3 (1%)	16	47
1	K	254/275 (92%)	244 (96%)	9 (4%)	1 (0%)	39	74
1	M	254/275 (92%)	242 (95%)	10 (4%)	2 (1%)	24	58
1	O	254/275 (92%)	244 (96%)	8 (3%)	2 (1%)	24	58
1	Q	254/275 (92%)	245 (96%)	7 (3%)	2 (1%)	24	58
1	S	254/275 (92%)	247 (97%)	6 (2%)	1 (0%)	39	74
1	U	254/275 (92%)	244 (96%)	8 (3%)	2 (1%)	24	58
1	W	254/275 (92%)	245 (96%)	8 (3%)	1 (0%)	39	74
2	B	239/248 (96%)	228 (95%)	10 (4%)	1 (0%)	39	74
2	D	245/248 (99%)	231 (94%)	12 (5%)	2 (1%)	24	58
2	F	239/248 (96%)	227 (95%)	10 (4%)	2 (1%)	24	58
2	H	239/248 (96%)	226 (95%)	12 (5%)	1 (0%)	39	74
2	J	239/248 (96%)	227 (95%)	11 (5%)	1 (0%)	39	74
2	L	245/248 (99%)	230 (94%)	13 (5%)	2 (1%)	24	58
2	N	239/248 (96%)	229 (96%)	9 (4%)	1 (0%)	39	74
2	P	246/248 (99%)	232 (94%)	12 (5%)	2 (1%)	24	58
2	R	243/248 (98%)	228 (94%)	11 (4%)	4 (2%)	12	38
2	T	245/248 (99%)	235 (96%)	9 (4%)	1 (0%)	39	74
2	V	242/248 (98%)	228 (94%)	13 (5%)	1 (0%)	39	74
2	X	239/248 (96%)	227 (95%)	11 (5%)	1 (0%)	39	74
All	All	5948/6276 (95%)	5679 (96%)	230 (4%)	39 (1%)	26	62

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	7	VAL
1	O	8	GLN
1	U	76	GLU
1	A	105	GLU
1	C	105	GLU

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	217/242 (90%)	204 (94%)	13 (6%)	24	56
1	C	219/242 (90%)	207 (94%)	12 (6%)	27	59
1	E	218/242 (90%)	208 (95%)	10 (5%)	33	67
1	G	217/242 (90%)	205 (94%)	12 (6%)	27	59
1	I	216/242 (89%)	205 (95%)	11 (5%)	29	63
1	K	217/242 (90%)	207 (95%)	10 (5%)	33	67
1	M	220/242 (91%)	206 (94%)	14 (6%)	22	52
1	O	218/242 (90%)	208 (95%)	10 (5%)	33	67
1	Q	218/242 (90%)	206 (94%)	12 (6%)	27	59
1	S	218/242 (90%)	206 (94%)	12 (6%)	27	59
1	U	218/242 (90%)	204 (94%)	14 (6%)	22	52
1	W	220/242 (91%)	207 (94%)	13 (6%)	24	57
2	B	198/208 (95%)	190 (96%)	8 (4%)	38	73
2	D	202/208 (97%)	197 (98%)	5 (2%)	55	86
2	F	196/208 (94%)	190 (97%)	6 (3%)	47	81
2	H	198/208 (95%)	190 (96%)	8 (4%)	38	73
2	J	198/208 (95%)	190 (96%)	8 (4%)	38	73
2	L	203/208 (98%)	193 (95%)	10 (5%)	31	65
2	N	198/208 (95%)	189 (96%)	9 (4%)	34	68
2	P	202/208 (97%)	192 (95%)	10 (5%)	30	64
2	R	200/208 (96%)	194 (97%)	6 (3%)	48	82
2	T	202/208 (97%)	194 (96%)	8 (4%)	38	73
2	V	200/208 (96%)	191 (96%)	9 (4%)	34	68
2	X	198/208 (95%)	191 (96%)	7 (4%)	43	77
All	All	5011/5400 (93%)	4774 (95%)	237 (5%)	32	67

5 of 237 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	164	VAL
2	N	208	SER
1	W	18	SER
2	L	233	ARG
1	M	191	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	178	ASN
2	N	178	ASN
2	T	40	ASN
2	L	6	GLN
2	L	40	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 39 ligands modelled in this entry, 13 are monoatomic - leaving 26 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	U5P	B	401	4	16,22,22	1.28	2 (12%)	21,33,33	3.12	5 (23%)
4	U5P	B	402	4	11,21,22	1.68	2 (18%)	15,30,33	4.09	6 (40%)
4	U5P	B	403	4	11,21,22	1.55	2 (18%)	15,30,33	4.11	5 (33%)
4	U5P	B	404	4	0,3,22	0.00	-	0,3,33	0.00	-
4	U5P	F	401	4	16,22,22	1.37	2 (12%)	21,33,33	3.25	5 (23%)
4	U5P	F	402	4	11,21,22	1.39	2 (18%)	15,30,33	3.77	2 (13%)
4	U5P	F	403	4	8,12,22	0.94	0	12,16,33	1.04	1 (8%)
4	U5P	F	404	4	8,12,22	0.61	0	12,16,33	0.84	0
4	U5P	H	401	4	13,13,22	0.97	0	18,19,33	1.24	2 (11%)
4	U5P	H	402	4	11,21,22	1.48	2 (18%)	15,30,33	3.65	2 (13%)
4	U5P	H	403	4	11,21,22	1.48	2 (18%)	15,30,33	3.45	3 (20%)
4	U5P	J	401	4	16,22,22	1.40	2 (12%)	21,33,33	3.26	4 (19%)
4	U5P	J	402	4	11,21,22	1.51	2 (18%)	15,30,33	3.79	3 (20%)
4	U5P	J	403	4	0,3,22	0.00	-	0,3,33	0.00	-
4	U5P	N	401	4	16,22,22	1.44	2 (12%)	21,33,33	3.32	3 (14%)
4	U5P	N	402	4	8,12,22	0.39	0	12,16,33	1.10	1 (8%)
4	U5P	N	403	4	8,12,22	0.65	0	12,16,33	0.89	0
5	RP5	N	405	-	13,13,14	1.26	1 (7%)	18,19,21	1.07	1 (5%)
4	U5P	V	401	4	16,22,22	1.45	2 (12%)	21,33,33	3.27	3 (14%)
4	U5P	V	402	4	11,21,22	1.60	2 (18%)	15,30,33	3.77	2 (13%)
4	U5P	V	403	4,6	11,21,22	1.49	2 (18%)	15,30,33	3.53	2 (13%)
4	U5P	V	404	4,6	11,21,22	1.51	2 (18%)	15,30,33	3.73	4 (26%)
4	U5P	X	401	4	16,22,22	1.35	2 (12%)	21,33,33	3.23	4 (19%)
4	U5P	X	402	4	11,21,22	1.45	2 (18%)	15,30,33	3.87	5 (33%)
4	U5P	X	403	4	8,12,22	0.90	0	12,16,33	1.17	2 (16%)
4	U5P	X	404	4	0,3,22	0.00	-	0,3,33	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	U5P	B	401	4	-	0/6/26/26	0/2/2/2
4	U5P	B	402	4	-	0/3/25/26	0/2/2/2
4	U5P	B	403	4	-	0/3/25/26	0/2/2/2
4	U5P	B	404	4	-	0/0/0/26	0/0/0/2
4	U5P	F	401	4	-	0/6/26/26	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	U5P	F	402	4	-	0/3/25/26	0/2/2/2
4	U5P	F	403	4	-	0/3/18/26	0/1/1/2
4	U5P	F	404	4	-	0/3/18/26	0/1/1/2
4	U5P	H	401	4	-	0/6/19/26	0/1/1/2
4	U5P	H	402	4	-	0/3/25/26	0/2/2/2
4	U5P	H	403	4	-	0/3/25/26	0/2/2/2
4	U5P	J	401	4	-	0/6/26/26	0/2/2/2
4	U5P	J	402	4	-	0/3/25/26	0/2/2/2
4	U5P	J	403	4	-	0/0/0/26	0/0/0/2
4	U5P	N	401	4	-	0/6/26/26	0/2/2/2
4	U5P	N	402	4	-	0/3/18/26	0/1/1/2
4	U5P	N	403	4	-	0/3/18/26	0/1/1/2
5	RP5	N	405	-	-	0/6/19/22	0/1/1/1
4	U5P	V	401	4	-	0/6/26/26	0/2/2/2
4	U5P	V	402	4	-	0/3/25/26	0/2/2/2
4	U5P	V	403	4,6	-	0/3/25/26	0/2/2/2
4	U5P	V	404	4,6	-	0/3/25/26	0/2/2/2
4	U5P	X	401	4	-	0/6/26/26	0/2/2/2
4	U5P	X	402	4	-	0/3/25/26	0/2/2/2
4	U5P	X	403	4	-	0/3/18/26	0/1/1/2
4	U5P	X	404	4	-	0/0/0/26	0/0/0/2

The worst 5 of 33 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	X	402	U5P	C6-N1	2.34	1.39	1.35
4	B	401	U5P	C6-N1	2.39	1.39	1.35
4	F	401	U5P	C6-N1	2.53	1.39	1.35
4	F	402	U5P	C6-N1	2.57	1.39	1.35
4	X	401	U5P	C6-N1	2.57	1.39	1.35

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	403	U5P	C4'-O4'-C1'	-5.61	103.55	109.72
4	F	401	U5P	C5-C4-N3	-3.61	113.85	123.12
4	V	402	U5P	C5-C4-N3	-3.57	113.97	123.12
4	B	402	U5P	O4'-C4'-C3'	-3.53	98.04	105.15
4	X	402	U5P	C5-C4-N3	-3.42	114.35	123.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	403	U5P	1	0
4	F	402	U5P	1	0
4	H	401	U5P	1	0
4	H	402	U5P	1	0
4	N	401	U5P	2	0
4	N	402	U5P	2	0
4	N	403	U5P	3	0
4	V	402	U5P	1	0
4	V	404	U5P	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	260/275 (94%)	0.62	21 (8%) 15 7	42, 50, 64, 71	0
1	C	260/275 (94%)	0.43	5 (1%) 70 59	42, 50, 64, 71	0
1	E	260/275 (94%)	0.46	12 (4%) 36 25	42, 50, 64, 71	0
1	G	260/275 (94%)	0.68	18 (6%) 20 11	42, 50, 64, 72	0
1	I	260/275 (94%)	0.50	12 (4%) 36 25	42, 50, 64, 71	0
1	K	260/275 (94%)	0.44	10 (3%) 44 32	42, 50, 64, 71	0
1	M	260/275 (94%)	0.53	6 (2%) 64 52	42, 50, 64, 72	0
1	O	260/275 (94%)	0.48	6 (2%) 64 52	42, 50, 64, 71	0
1	Q	260/275 (94%)	0.53	6 (2%) 64 52	42, 50, 64, 71	0
1	S	260/275 (94%)	0.46	5 (1%) 70 59	42, 50, 64, 71	0
1	U	260/275 (94%)	0.55	13 (5%) 32 21	42, 50, 64, 72	0
1	W	260/275 (94%)	0.45	8 (3%) 52 40	42, 50, 64, 71	0
2	B	241/248 (97%)	0.55	10 (4%) 41 29	42, 49, 65, 78	0
2	D	247/248 (99%)	0.46	7 (2%) 56 44	42, 49, 65, 78	0
2	F	241/248 (97%)	0.49	6 (2%) 61 48	42, 49, 65, 78	0
2	H	241/248 (97%)	0.56	13 (5%) 29 19	42, 49, 65, 78	0
2	J	241/248 (97%)	0.56	11 (4%) 36 25	42, 49, 65, 78	0
2	L	247/248 (99%)	0.54	7 (2%) 56 44	42, 49, 65, 78	0
2	N	241/248 (97%)	0.53	7 (2%) 55 43	42, 49, 65, 78	0
2	P	248/248 (100%)	0.52	11 (4%) 38 26	42, 49, 65, 78	0
2	R	245/248 (98%)	0.52	4 (1%) 74 66	42, 49, 65, 78	0
2	T	247/248 (99%)	0.44	1 (0%) 93 90	42, 49, 65, 78	0
2	V	244/248 (98%)	0.60	8 (3%) 50 38	42, 49, 65, 78	0
2	X	241/248 (97%)	0.48	8 (3%) 50 38	42, 49, 65, 78	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å <sup>2</sup> )	Q<0.9
All	All	6044/6276 (96%)	0.52	215 (3%)	46	34	42, 49, 65, 78	0

The worst 5 of 215 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	275	ILE	6.6
2	F	248	VAL	6.6
1	U	275	ILE	6.0
2	D	248	VAL	6.0
2	T	248	VAL	5.6

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CL	L	301	1/1	0.95	0.33	3.92	76,76,76,76	0
4	U5P	V	404	20/21	0.83	0.28	0.92	85,88,89,89	0
3	CL	N	301	1/1	0.92	0.23	0.62	59,59,59,59	0
5	RP5	N	405	13/14	0.90	0.23	0.54	66,67,69,69	0
3	CL	J	301	1/1	0.94	0.22	0.44	50,50,50,50	0
4	U5P	F	404	12/21	0.81	0.25	0.37	109,110,111,111	0
3	CL	H	301	1/1	0.81	0.22	0.36	66,66,66,66	0
3	CL	V	301	1/1	0.93	0.22	0.11	45,45,45,45	0
3	CL	X	301	1/1	0.91	0.20	0.01	52,52,52,52	0
4	U5P	F	401	21/21	0.80	0.21	-0.04	107,108,110,110	0
4	U5P	V	401	21/21	0.85	0.23	-0.20	78,81,82,83	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	U5P	B	401	21/21	0.79	0.23	-0.25	94,94,94,96	0
4	U5P	J	401	21/21	0.78	0.23	-0.33	114,115,117,117	0
4	U5P	X	401	21/21	0.88	0.22	-0.48	90,92,95,95	0
3	CL	P	301	1/1	0.88	0.20	-0.80	80,80,80,80	0
4	U5P	B	402	20/21	0.79	0.18	-1.20	95,98,99,100	0
4	U5P	N	401	21/21	0.80	0.20	-1.32	90,93,94,95	0
4	U5P	H	403	20/21	0.81	0.19	-1.43	101,106,107,107	0
3	CL	D	301	1/1	0.93	0.11	-1.70	76,76,76,76	0
3	CL	F	301	1/1	0.98	0.12	-2.11	36,36,36,36	0
4	U5P	H	401	13/21	0.81	0.19	-2.13	100,102,105,105	0
3	CL	R	301	1/1	0.79	0.15	-2.42	63,63,63,63	0
3	CL	T	301	1/1	0.88	0.13	-3.15	57,57,57,57	0
3	CL	B	301	1/1	0.94	0.11	-3.41	55,55,55,55	0
4	U5P	X	402	20/21	0.88	0.21	-	95,98,99,99	0
4	U5P	X	403	12/21	0.78	0.23	-	97,97,98,98	0
4	U5P	N	403	12/21	0.82	0.22	-	84,85,88,89	0
4	U5P	F	402	20/21	0.81	0.24	-	105,106,107,108	0
4	U5P	J	402	20/21	0.73	0.25	-	107,108,112,113	0
4	U5P	X	404	4/21	0.78	0.23	-	97,97,98,98	0
4	U5P	J	403	4/21	0.89	0.16	-	106,106,106,107	0
6	NA	V	501	1/1	0.97	0.13	-	40,40,40,40	0
4	U5P	F	403	12/21	0.82	0.20	-	107,108,109,110	0
4	U5P	V	403	20/21	0.86	0.17	-	74,83,84,85	0
4	U5P	V	402	20/21	0.89	0.20	-	73,74,75,76	0
4	U5P	H	402	20/21	0.79	0.20	-	98,100,101,101	0
4	U5P	B	404	4/21	0.87	0.22	-	101,101,102,102	0
4	U5P	N	402	12/21	0.89	0.16	-	82,86,89,89	0
4	U5P	B	403	20/21	0.78	0.23	-	100,105,107,107	0

## 6.5 Other polymers

There are no such residues in this entry.