



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:28 pm GMT

PDB ID : 2C38  
Title : RNASE PH CORE OF THE ARCHAEAL EXOSOME IN COMPLEX WITH A5 RNA  
Authors : Lorentzen, E.; Conti, E.  
Deposited on : 2005-10-04  
Resolution : 3.10 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : trunk28620

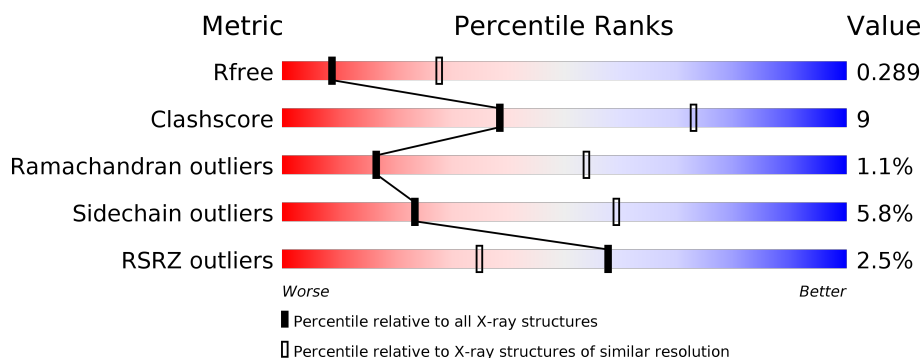
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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>2%</div> <div> <div></div> <div>81%</div> <div>17%</div> <div>..</div> </div> </div>
1	C	275	<div> <div>2%</div> <div> <div></div> <div>77%</div> <div>20%</div> <div>..</div> </div> </div>
1	E	275	<div> <div>5%</div> <div> <div></div> <div>79%</div> <div>16%</div> <div>..</div> </div> </div>
1	G	275	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>21%</div> <div>..</div> </div> </div>
1	I	275	<div> <div>4%</div> <div> <div></div> <div>75%</div> <div>21%</div> <div>..</div> </div> </div>
1	K	275	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>21%</div> <div>..</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
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	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 47618 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	271	Total	C	N	O	S	0	0	0
			2032	1299	334	394	5			
1	C	271	Total	C	N	O	S	0	0	0
			2044	1306	336	397	5			
1	E	271	Total	C	N	O	S	0	0	0
			2042	1305	335	397	5			
1	G	271	Total	C	N	O	S	0	0	0
			2038	1303	335	395	5			
1	I	271	Total	C	N	O	S	0	0	0
			2034	1300	334	395	5			
1	K	271	Total	C	N	O	S	0	0	0
			2042	1305	335	397	5			
1	M	271	Total	C	N	O	S	0	0	0
			2046	1306	334	401	5			
1	O	271	Total	C	N	O	S	0	0	0
			2040	1304	335	396	5			
1	Q	271	Total	C	N	O	S	0	0	0
			2045	1307	335	398	5			
1	S	271	Total	C	N	O	S	0	0	0
			2049	1309	335	400	5			
1	U	270	Total	C	N	O	S	0	0	0
			2043	1304	335	399	5			
1	W	271	Total	C	N	O	S	0	0	0
			2044	1306	335	398	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
			1855	1170	321	354	10			
2	D	247	Total	C	N	O	S	0	0	0
			1893	1195	328	359	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
			1859	1172	321	356	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
			1855	1170	321	354	10			
2	L	247	Total	C	N	O	S	0	0	0
			1901	1199	328	363	11			
2	N	241	Total	C	N	O	S	0	0	0
			1849	1166	317	356	10			
2	P	248	Total	C	N	O	S	0	0	0
			1909	1204	329	364	12			
2	R	247	Total	C	N	O	S	0	0	0
			1902	1198	331	363	10			
2	T	247	Total	C	N	O	S	0	0	0
			1901	1199	328	363	11			
2	V	239	Total	C	N	O	S	0	0	0
			1842	1162	319	351	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).

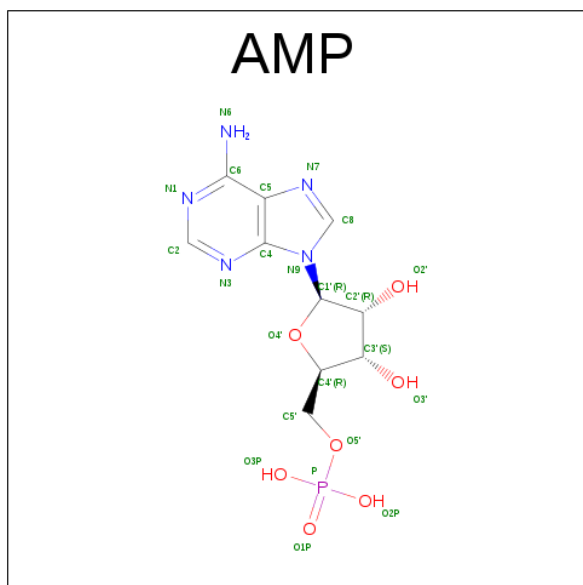
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total	Cl	0	0
			1	1		
3	J	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		
3	H	1	Total	Cl	0	0
			1	1		
3	B	1	Total	Cl	0	0
			1	1		
3	V	1	Total	Cl	0	0
			1	1		
3	T	1	Total	Cl	0	0
			1	1		
3	N	1	Total	Cl	0	0
			1	1		
3	X	1	Total	Cl	0	0
			1	1		

*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 ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula:  $C_{10}H_{14}N_5O_7P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
4	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	F	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
4	F	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	F	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	F	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

Continued on next page...

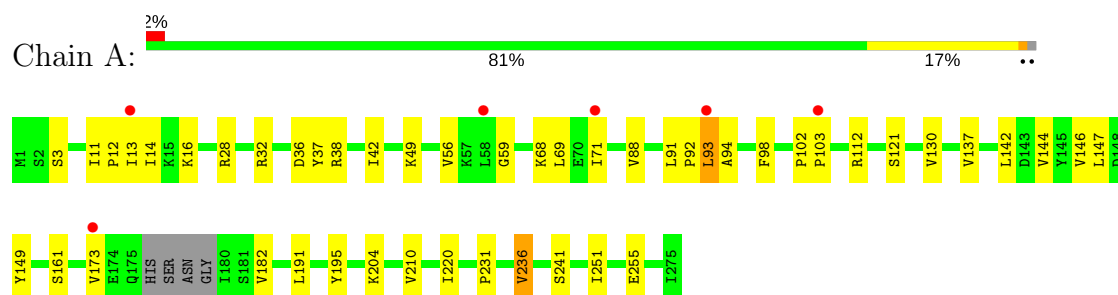
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	H	1	Total 23	C 10	N 5	O 7	P 1	0	0
4	H	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	H	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	H	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	J	1	Total 23	C 10	N 5	O 7	P 1	0	0
4	J	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	J	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	J	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	N	1	Total 23	C 10	N 5	O 7	P 1	0	0
4	N	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	N	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	N	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	V	1	Total 23	C 10	N 5	O 7	P 1	0	0
4	V	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	V	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	V	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	X	1	Total 23	C 10	N 5	O 7	P 1	0	0
4	X	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	X	1	Total 22	C 10	N 5	O 6	P 1	0	0
4	X	1	Total 22	C 10	N 5	O 6	P 1	0	0

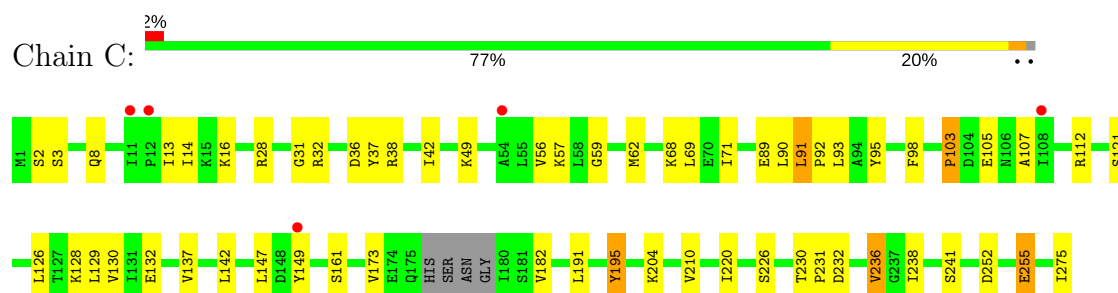
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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 ( $RSRZ > 2$ ). 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.

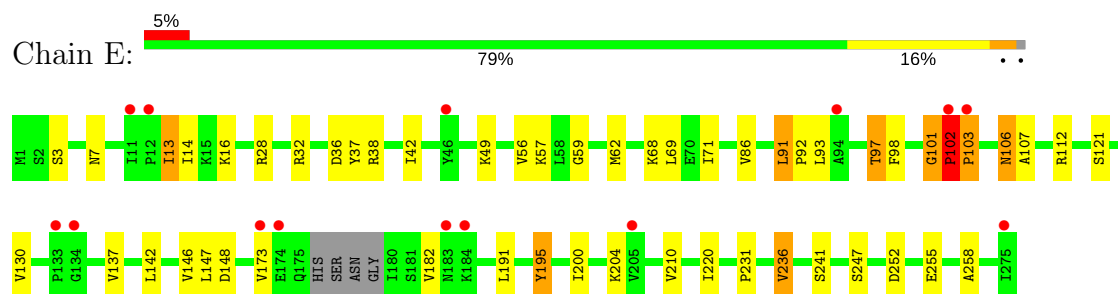
#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



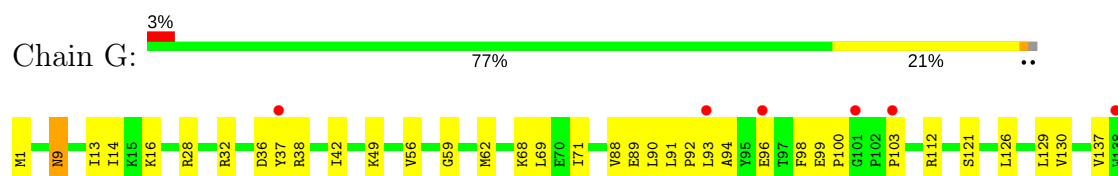
#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



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



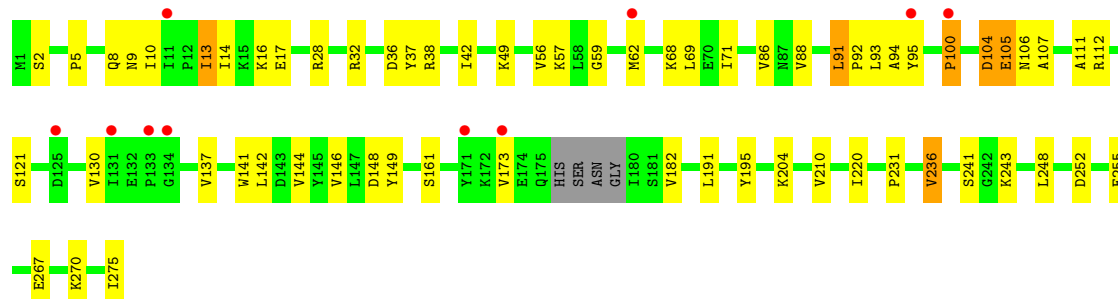
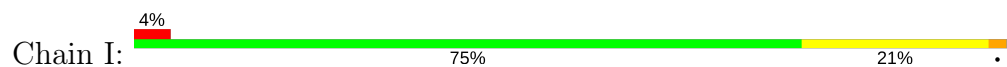
#### • Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



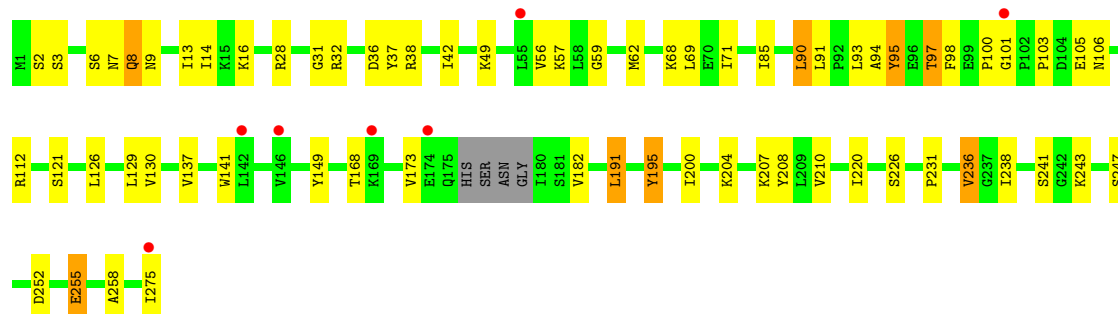




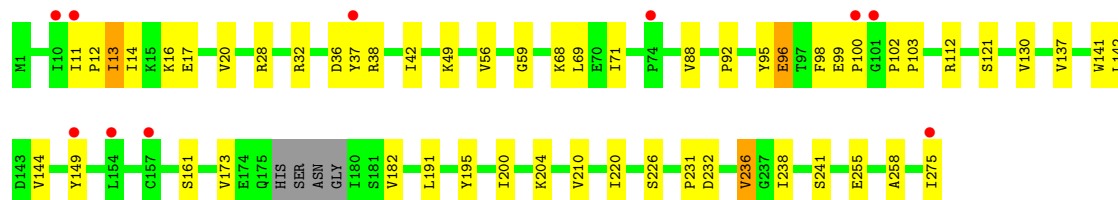
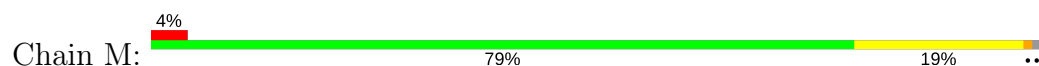
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



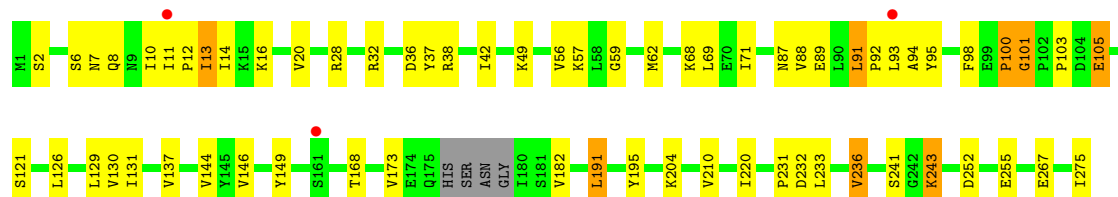
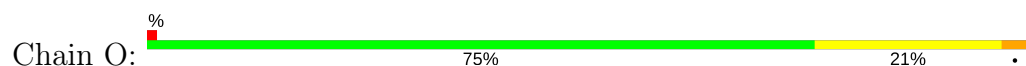
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



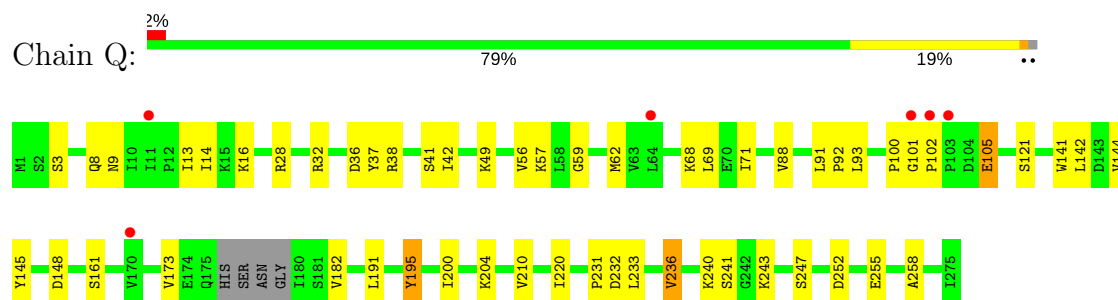
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



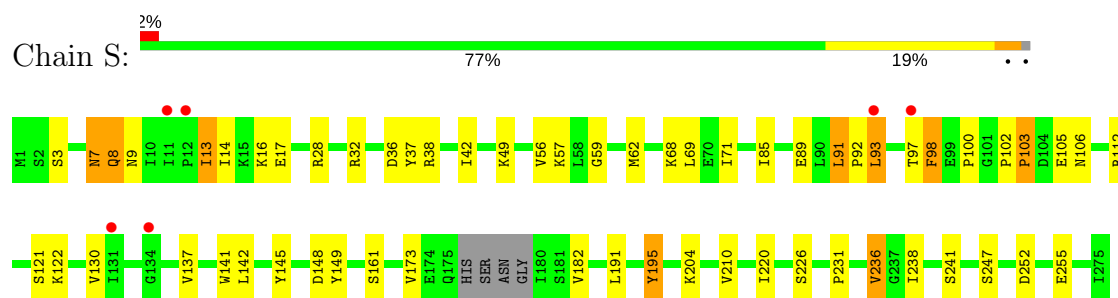
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



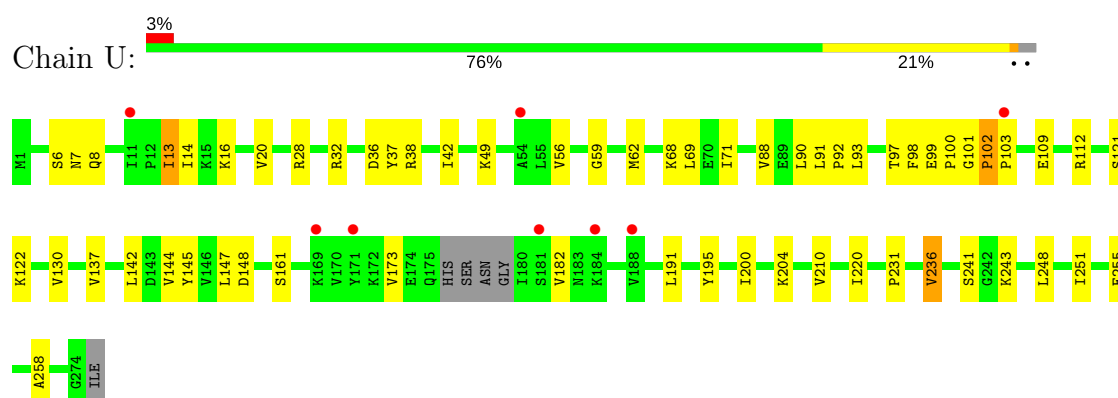
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



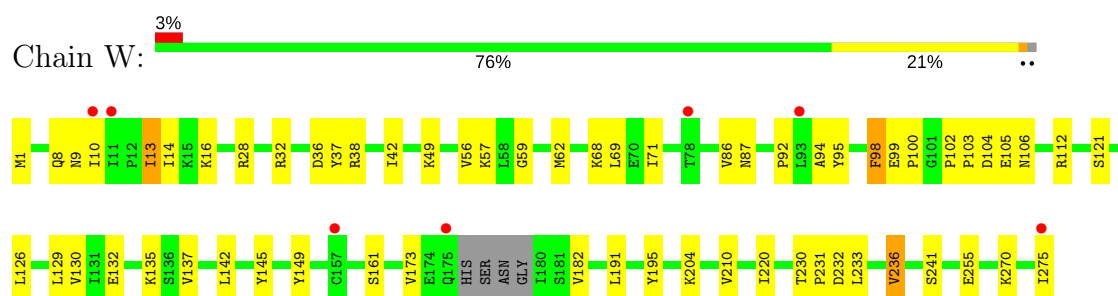
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2



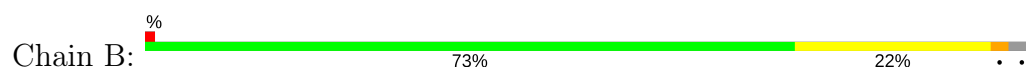
• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2

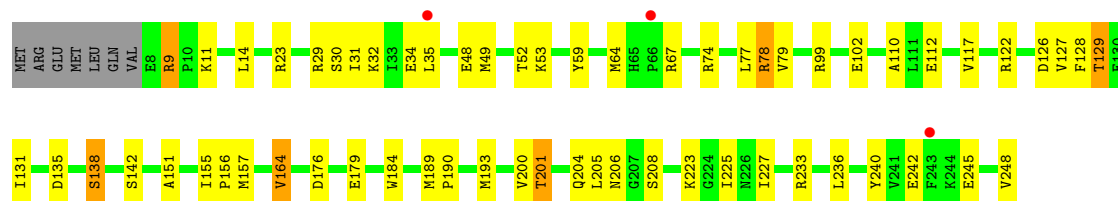


• Molecule 1: PROBABLE EXOSOME COMPLEX EXONUCLEASE 2

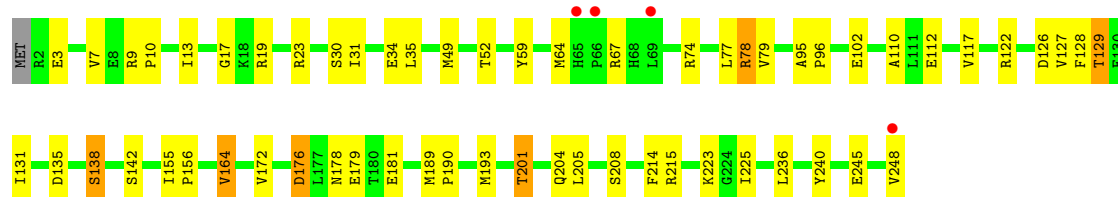
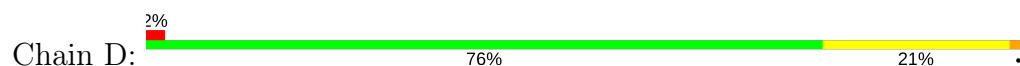


• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1

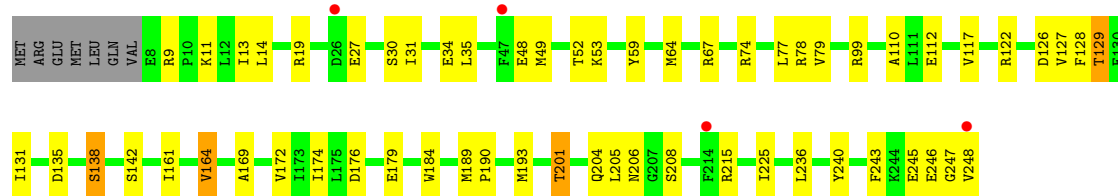
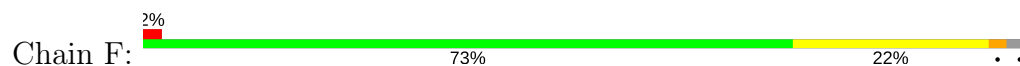




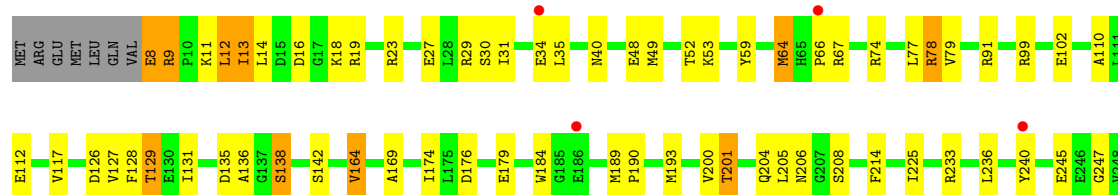
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



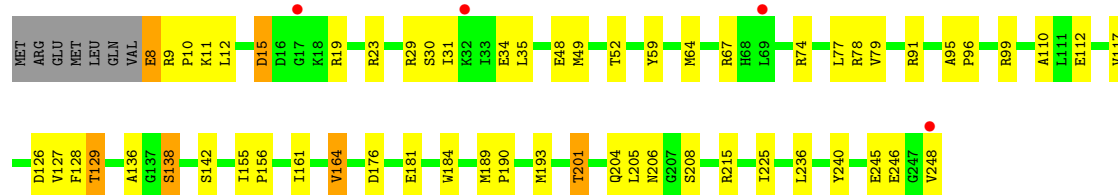
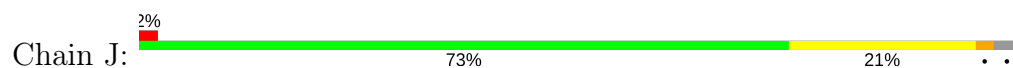
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



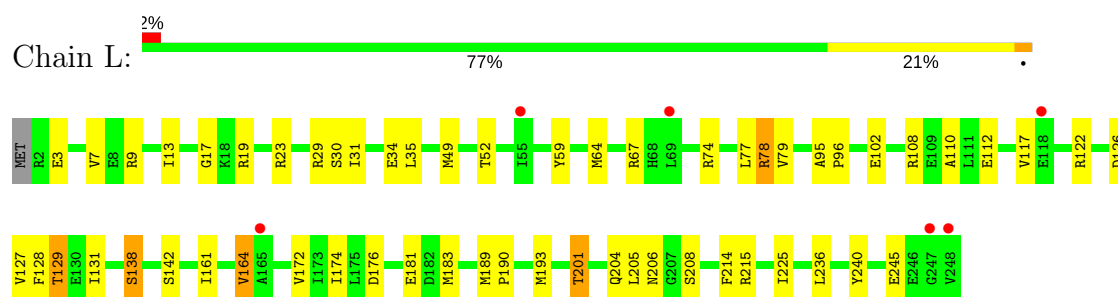
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



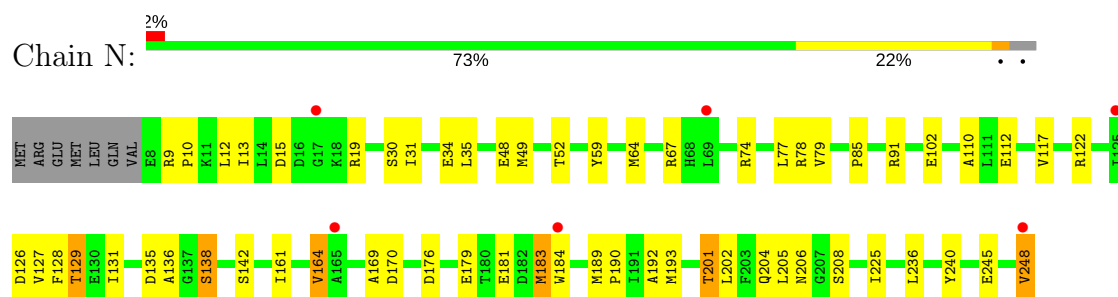
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



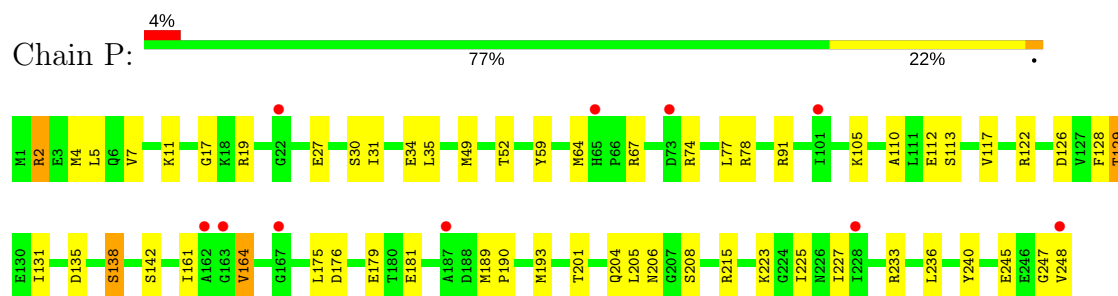
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



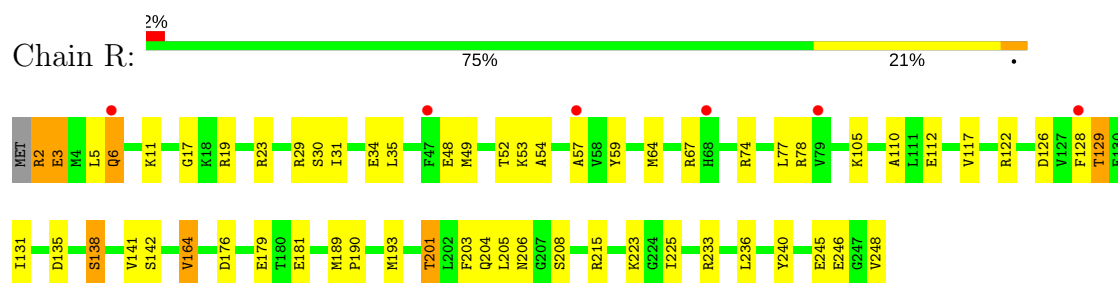
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



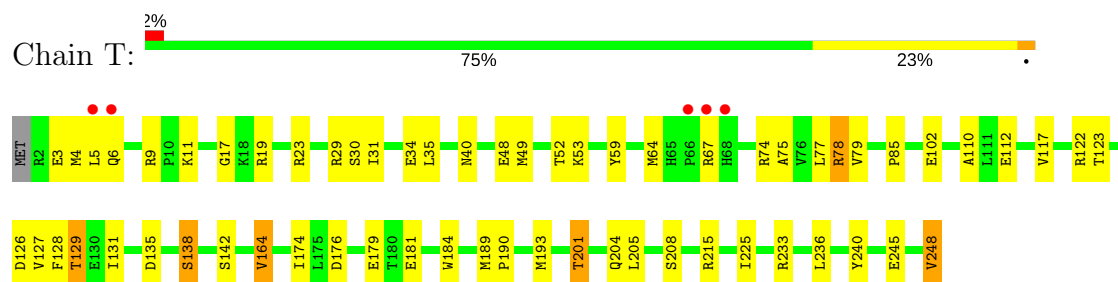
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



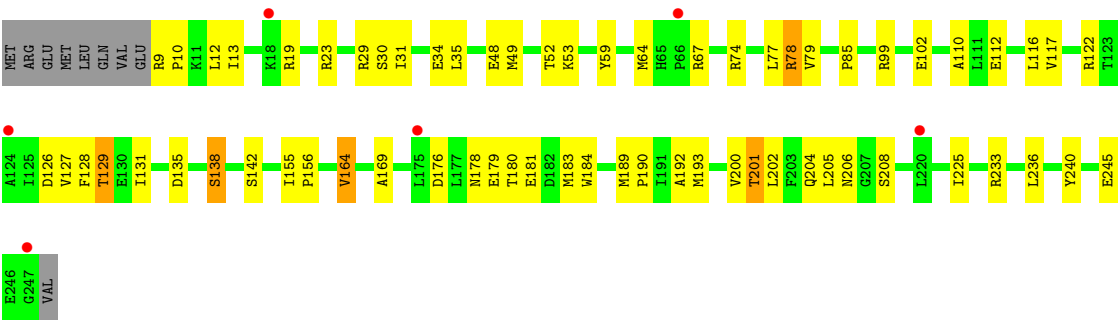
• Molecule 2: PROBABLE EXOSOME COMPLEX EXONUCLEASE 1



• 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, $\alpha$ , $\beta$ , $\gamma$	206.26Å 213.61Å 434.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.78 – 3.10 93.15 – 3.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (95.78-3.10) 91.8 (93.15-3.10)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 3.13Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.276 , 0.289 0.272 , 0.289	Depositor DCC
$R_{free}$ test set	5102 reflections (3.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.4	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 12.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	0.100 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	47618	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.89 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.2640e-03.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AMP, 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	0.46	0/2063	0.58	1/2809 (0.0%)
1	C	0.50	0/2075	0.60	0/2823
1	E	0.46	0/2073	0.56	0/2821
1	G	0.48	0/2069	0.60	0/2816
1	I	0.49	0/2065	0.58	0/2810
1	K	0.49	0/2073	0.60	0/2821
1	M	0.46	0/2077	0.59	0/2827
1	O	0.49	0/2071	0.60	0/2818
1	Q	0.49	0/2076	0.59	0/2824
1	S	0.48	0/2080	0.61	1/2831 (0.0%)
1	U	0.47	0/2074	0.56	0/2822
1	W	0.49	0/2075	0.60	0/2823
2	B	0.43	0/1883	0.61	0/2544
2	D	0.46	0/1921	0.61	0/2596
2	F	0.44	0/1887	0.60	0/2549
2	H	0.45	0/1887	0.60	0/2549
2	J	0.44	0/1883	0.61	0/2544
2	L	0.48	0/1929	0.62	0/2606
2	N	0.44	0/1877	0.60	0/2538
2	P	0.46	0/1937	0.62	0/2616
2	R	0.47	0/1930	0.61	0/2607
2	T	0.49	0/1929	0.64	1/2606 (0.0%)
2	V	0.42	0/1870	0.60	0/2527
2	X	0.44	0/1887	0.59	0/2549
All	All	0.47	0/47691	0.60	3/64676 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	Q	0	1
2	D	0	1
2	L	0	1
2	P	0	1
2	R	0	1
2	T	0	1
All	All	0	6

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	5	LEU	CA-CB-CG	6.04	129.19	115.30
1	S	93	LEU	CA-CB-CG	5.29	127.45	115.30
1	A	93	LEU	CA-CB-CG	5.03	126.86	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	17	GLY	Peptide
2	L	17	GLY	Peptide
2	P	17	GLY	Peptide
1	Q	102	PRO	Peptide
2	R	17	GLY	Peptide

## 5.2 Too-close contacts [i](#)

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	2032	0	2057	26	0
1	C	2044	0	2078	36	0
1	E	2042	0	2077	38	0
1	G	2038	0	2073	36	0
1	I	2034	0	2059	40	0
1	K	2042	0	2077	42	0
1	M	2046	0	2074	27	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	O	2040	0	2072	34	0
1	Q	2045	0	2078	25	0
1	S	2049	0	2085	34	0
1	U	2043	0	2075	33	0
1	W	2044	0	2076	33	0
2	B	1855	0	1880	43	0
2	D	1893	0	1917	36	0
2	F	1859	0	1884	38	0
2	H	1859	0	1884	43	0
2	J	1855	0	1880	37	0
2	L	1901	0	1925	34	0
2	N	1849	0	1862	36	0
2	P	1909	0	1937	34	0
2	R	1902	0	1922	39	0
2	T	1901	0	1925	39	0
2	V	1842	0	1869	46	0
2	X	1859	0	1884	43	0
3	B	1	0	0	1	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	1	0
3	L	1	0	0	0	0
3	N	1	0	0	1	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	1	0
3	X	1	0	0	0	0
4	B	89	0	45	4	0
4	F	89	0	45	5	0
4	H	89	0	45	4	0
4	J	89	0	45	6	0
4	N	89	0	45	3	0
4	V	89	0	45	6	0
4	X	89	0	45	1	0
All	All	47618	0	47965	818	0

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 9.

The worst 5 of 818 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)
2:D:223:LYS:HB3	2:D:248:VAL:HG11	1.41	1.00
2:X:9:ARG:HG2	2:X:184:TRP:CE3	2.01	0.95
2:B:223:LYS:HB3	2:B:248:VAL:HG11	1.52	0.90
2:F:77:LEU:HD12	2:F:112:GLU:HG2	1.53	0.90
2:B:77:LEU:HD12	2:B:112:GLU:HG2	1.56	0.88

There are no symmetry-related clashes.

## 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	267/275 (97%)	247 (92%)	18 (7%)	2 (1%)	25	64
1	C	267/275 (97%)	247 (92%)	19 (7%)	1 (0%)	38	75
1	E	267/275 (97%)	242 (91%)	22 (8%)	3 (1%)	17	54
1	G	267/275 (97%)	247 (92%)	17 (6%)	3 (1%)	17	54
1	I	267/275 (97%)	243 (91%)	18 (7%)	6 (2%)	8	35
1	K	267/275 (97%)	246 (92%)	15 (6%)	6 (2%)	8	35
1	M	267/275 (97%)	248 (93%)	17 (6%)	2 (1%)	25	64
1	O	267/275 (97%)	247 (92%)	13 (5%)	7 (3%)	6	31
1	Q	267/275 (97%)	249 (93%)	14 (5%)	4 (2%)	12	45
1	S	267/275 (97%)	244 (91%)	18 (7%)	5 (2%)	9	39
1	U	266/275 (97%)	248 (93%)	15 (6%)	3 (1%)	17	54
1	W	267/275 (97%)	242 (91%)	19 (7%)	6 (2%)	8	35
2	B	239/248 (96%)	224 (94%)	14 (6%)	1 (0%)	38	75
2	D	245/248 (99%)	230 (94%)	13 (5%)	2 (1%)	22	62
2	F	239/248 (96%)	226 (95%)	12 (5%)	1 (0%)	38	75
2	H	239/248 (96%)	226 (95%)	12 (5%)	1 (0%)	38	75

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	J	239/248 (96%)	221 (92%)	15 (6%)	3 (1%)	14	48
2	L	245/248 (99%)	227 (93%)	17 (7%)	1 (0%)	38	75
2	N	239/248 (96%)	226 (95%)	12 (5%)	1 (0%)	38	75
2	P	246/248 (99%)	228 (93%)	17 (7%)	1 (0%)	38	75
2	R	245/248 (99%)	227 (93%)	15 (6%)	3 (1%)	15	51
2	T	245/248 (99%)	229 (94%)	14 (6%)	2 (1%)	22	62
2	V	237/248 (96%)	223 (94%)	12 (5%)	2 (1%)	22	62
2	X	239/248 (96%)	225 (94%)	12 (5%)	2 (1%)	22	62
All	All	6100/6276 (97%)	5662 (93%)	370 (6%)	68 (1%)	17	54

5 of 68 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103	PRO
1	E	102	PRO
1	I	10	ILE
1	I	93	LEU
1	I	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	221/242 (91%)	212 (96%)	9 (4%)	35	72
1	C	224/242 (93%)	209 (93%)	15 (7%)	19	54
1	E	224/242 (93%)	208 (93%)	16 (7%)	17	52
1	G	223/242 (92%)	212 (95%)	11 (5%)	29	66
1	I	221/242 (91%)	208 (94%)	13 (6%)	23	58
1	K	224/242 (93%)	208 (93%)	16 (7%)	17	52
1	M	225/242 (93%)	213 (95%)	12 (5%)	26	63
1	O	223/242 (92%)	207 (93%)	16 (7%)	17	51

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	Q	224/242 (93%)	209 (93%)	15 (7%)	19	54
1	S	226/242 (93%)	208 (92%)	18 (8%)	14	47
1	U	225/242 (93%)	214 (95%)	11 (5%)	29	66
1	W	224/242 (93%)	214 (96%)	10 (4%)	32	68
2	B	197/208 (95%)	186 (94%)	11 (6%)	25	61
2	D	200/208 (96%)	191 (96%)	9 (4%)	32	68
2	F	198/208 (95%)	189 (96%)	9 (4%)	32	68
2	H	198/208 (95%)	183 (92%)	15 (8%)	15	49
2	J	197/208 (95%)	186 (94%)	11 (6%)	25	61
2	L	202/208 (97%)	189 (94%)	13 (6%)	20	56
2	N	196/208 (94%)	185 (94%)	11 (6%)	25	61
2	P	203/208 (98%)	191 (94%)	12 (6%)	23	58
2	R	201/208 (97%)	192 (96%)	9 (4%)	32	68
2	T	202/208 (97%)	191 (95%)	11 (5%)	26	62
2	V	196/208 (94%)	186 (95%)	10 (5%)	28	64
2	X	198/208 (95%)	187 (94%)	11 (6%)	25	61
All	All	5072/5400 (94%)	4778 (94%)	294 (6%)	23	59

5 of 294 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	247	SER
2	N	138	SER
2	V	208	SER
2	L	7	VAL
1	M	17	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	206	ASN
1	O	87	ASN
2	V	206	ASN
2	N	178	ASN
2	N	206	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 40 ligands modelled in this entry, 12 are monoatomic - leaving 28 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	AMP	B	401	4	22,25,25	0.77	0	24,38,38	1.39	3 (12%)
4	AMP	B	402	4	18,24,25	0.63	0	17,35,38	0.78	0
4	AMP	B	403	4	18,24,25	0.62	0	17,35,38	1.03	1 (5%)
4	AMP	B	404	4	18,24,25	0.75	1 (5%)	17,35,38	0.94	1 (5%)
4	AMP	F	401	4	22,25,25	0.63	0	24,38,38	1.33	4 (16%)
4	AMP	F	402	4	18,24,25	0.75	0	17,35,38	0.99	1 (5%)
4	AMP	F	403	4	18,24,25	0.69	0	17,35,38	0.72	0
4	AMP	F	404	4	18,24,25	0.70	0	17,35,38	1.00	1 (5%)
4	AMP	H	401	4	22,25,25	0.70	0	24,38,38	0.88	1 (4%)
4	AMP	H	402	4	18,24,25	0.64	0	17,35,38	0.85	1 (5%)
4	AMP	H	403	4	18,24,25	0.70	0	17,35,38	0.78	0
4	AMP	H	404	4	18,24,25	0.70	0	17,35,38	0.99	0
4	AMP	J	401	4	22,25,25	0.65	0	24,38,38	1.69	4 (16%)
4	AMP	J	402	4	18,24,25	0.65	0	17,35,38	0.78	0
4	AMP	J	403	4	18,24,25	0.70	0	17,35,38	0.73	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	AMP	J	404	4	18,24,25	0.70	0	17,35,38	1.05	1 (5%)
4	AMP	N	401	4	22,25,25	0.72	0	24,38,38	0.99	1 (4%)
4	AMP	N	402	4	18,24,25	0.61	0	17,35,38	0.66	0
4	AMP	N	403	4	18,24,25	0.79	1 (5%)	17,35,38	0.77	1 (5%)
4	AMP	N	404	4	18,24,25	0.85	1 (5%)	17,35,38	1.22	2 (11%)
4	AMP	V	401	4	22,25,25	0.71	0	24,38,38	0.83	0
4	AMP	V	402	4	18,24,25	0.77	0	17,35,38	0.96	1 (5%)
4	AMP	V	403	4	18,24,25	0.61	0	17,35,38	0.77	0
4	AMP	V	404	4	18,24,25	0.85	1 (5%)	17,35,38	1.40	2 (11%)
4	AMP	X	401	4	22,25,25	0.69	0	24,38,38	0.98	2 (8%)
4	AMP	X	402	4	18,24,25	0.74	0	17,35,38	1.12	1 (5%)
4	AMP	X	403	4	18,24,25	0.75	0	17,35,38	0.85	0
4	AMP	X	404	2,4	18,24,25	0.86	1 (5%)	17,35,38	0.85	0

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	AMP	B	401	4	-	0/6/26/26	0/3/3/3
4	AMP	B	402	4	-	0/3/25/26	0/3/3/3
4	AMP	B	403	4	-	0/3/25/26	0/3/3/3
4	AMP	B	404	4	-	0/3/25/26	0/3/3/3
4	AMP	F	401	4	-	0/6/26/26	0/3/3/3
4	AMP	F	402	4	-	0/3/25/26	0/3/3/3
4	AMP	F	403	4	-	0/3/25/26	0/3/3/3
4	AMP	F	404	4	-	0/3/25/26	0/3/3/3
4	AMP	H	401	4	-	0/6/26/26	0/3/3/3
4	AMP	H	402	4	-	0/3/25/26	0/3/3/3
4	AMP	H	403	4	-	0/3/25/26	0/3/3/3
4	AMP	H	404	4	-	0/3/25/26	0/3/3/3
4	AMP	J	401	4	-	0/6/26/26	0/3/3/3
4	AMP	J	402	4	-	0/3/25/26	0/3/3/3
4	AMP	J	403	4	-	0/3/25/26	0/3/3/3
4	AMP	J	404	4	-	0/3/25/26	0/3/3/3
4	AMP	N	401	4	-	0/6/26/26	0/3/3/3
4	AMP	N	402	4	-	0/3/25/26	0/3/3/3
4	AMP	N	403	4	-	0/3/25/26	0/3/3/3
4	AMP	N	404	4	-	0/3/25/26	0/3/3/3
4	AMP	V	401	4	-	0/6/26/26	0/3/3/3

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	AMP	V	402	4	-	0/3/25/26	0/3/3/3
4	AMP	V	403	4	-	0/3/25/26	0/3/3/3
4	AMP	V	404	4	-	0/3/25/26	0/3/3/3
4	AMP	X	401	4	-	0/6/26/26	0/3/3/3
4	AMP	X	402	4	-	0/3/25/26	0/3/3/3
4	AMP	X	403	4	-	0/3/25/26	0/3/3/3
4	AMP	X	404	2,4	-	0/3/25/26	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	X	404	AMP	O5'-C5'	-2.70	1.41	1.44
4	N	404	AMP	O5'-C5'	-2.44	1.41	1.44
4	V	404	AMP	O5'-C5'	-2.38	1.41	1.44
4	N	403	AMP	O5'-C5'	-2.27	1.41	1.44
4	B	404	AMP	O5'-C5'	-2.02	1.41	1.44

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	401	AMP	C4'-O4'-C1'	-4.55	104.93	109.77
4	J	401	AMP	C2'-C3'-C4'	-4.39	94.08	102.62
4	V	404	AMP	C4'-O4'-C1'	-4.18	105.32	109.77
4	F	401	AMP	C4'-O4'-C1'	-3.61	105.93	109.77
4	B	401	AMP	C2'-C3'-C4'	-3.13	96.52	102.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

14 monomers are involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	401	AMP	1	0
4	B	403	AMP	1	0
4	B	404	AMP	2	0
4	F	403	AMP	2	0
4	F	404	AMP	3	0
4	H	403	AMP	1	0
4	H	404	AMP	4	0
4	J	403	AMP	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	J	404	AMP	3	0
4	N	403	AMP	1	0
4	N	404	AMP	2	0
4	V	403	AMP	1	0
4	V	404	AMP	5	0
4	X	404	AMP	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	271/275 (98%)	0.51	6 (2%)	62 41	49, 56, 75, 83	0
1	C	271/275 (98%)	0.50	5 (1%)	69 47	49, 56, 71, 78	0
1	E	271/275 (98%)	0.52	14 (5%)	28 12	49, 56, 73, 80	0
1	G	271/275 (98%)	0.54	8 (2%)	51 27	49, 56, 76, 88	0
1	I	271/275 (98%)	0.53	10 (3%)	42 21	49, 56, 75, 82	0
1	K	271/275 (98%)	0.52	7 (2%)	56 33	49, 56, 73, 79	0
1	M	271/275 (98%)	0.50	10 (3%)	42 21	49, 56, 76, 83	0
1	O	271/275 (98%)	0.52	3 (1%)	80 65	49, 56, 75, 85	0
1	Q	271/275 (98%)	0.52	6 (2%)	62 41	49, 56, 74, 79	0
1	S	271/275 (98%)	0.46	6 (2%)	62 41	49, 56, 76, 84	0
1	U	270/275 (98%)	0.46	8 (2%)	51 27	49, 56, 75, 79	0
1	W	271/275 (98%)	0.49	7 (2%)	56 33	49, 56, 73, 78	0
2	B	241/248 (97%)	0.51	3 (1%)	79 61	49, 55, 71, 85	0
2	D	247/248 (99%)	0.55	4 (1%)	72 51	49, 55, 74, 93	0
2	F	241/248 (97%)	0.44	4 (1%)	70 49	49, 55, 71, 86	0
2	H	241/248 (97%)	0.53	4 (1%)	70 49	49, 55, 71, 85	0
2	J	241/248 (97%)	0.49	4 (1%)	70 49	49, 55, 71, 86	0
2	L	247/248 (99%)	0.52	6 (2%)	59 37	49, 55, 74, 85	0
2	N	241/248 (97%)	0.52	6 (2%)	58 35	49, 55, 71, 85	0
2	P	248/248 (100%)	0.56	10 (4%)	39 19	49, 55, 75, 84	0
2	R	247/248 (99%)	0.55	6 (2%)	59 37	49, 55, 74, 85	0
2	T	247/248 (99%)	0.55	5 (2%)	65 44	49, 55, 74, 89	0
2	V	239/248 (96%)	0.52	6 (2%)	58 35	49, 55, 71, 84	0
2	X	241/248 (97%)	0.50	5 (2%)	64 43	49, 55, 71, 84	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	6172/6276 (98%)	0.51	153 (2%) 58 35	49, 56, 73, 93	0

The worst 5 of 153 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	248	VAL	6.4
1	G	103	PRO	6.2
2	L	248	VAL	5.2
2	J	248	VAL	5.1
1	E	94	ALA	4.9

## 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
4	AMP	J	404	22/23	0.92	0.28	1.04	88,92,93,93	0
4	AMP	N	404	22/23	0.92	0.27	0.58	85,87,87,87	0
3	CL	T	301	1/1	0.99	0.28	0.28	10,10,10,10	0
4	AMP	F	404	22/23	0.91	0.24	0.27	88,93,94,95	0
4	AMP	X	404	22/23	0.93	0.29	0.17	88,89,89,89	0
4	AMP	H	404	22/23	0.91	0.25	-0.13	74,78,80,80	0
4	AMP	B	404	22/23	0.94	0.24	-0.22	86,91,93,93	0
4	AMP	X	403	22/23	0.91	0.24	-0.39	83,84,87,87	0
3	CL	J	301	1/1	0.98	0.25	-0.40	19,19,19,19	0
4	AMP	F	401	23/23	0.89	0.24	-0.43	81,85,90,91	0
3	CL	N	301	1/1	0.98	0.21	-0.58	12,12,12,12	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	AMP	V	404	22/23	0.94	0.23	-0.60	68,69,72,73	0
4	AMP	F	403	22/23	0.93	0.21	-0.72	85,90,91,91	0
4	AMP	N	403	22/23	0.94	0.20	-0.75	84,84,86,87	0
4	AMP	F	402	22/23	0.94	0.19	-0.80	80,82,84,85	0
4	AMP	V	401	23/23	0.90	0.23	-0.81	73,75,79,79	0
3	CL	F	301	1/1	0.99	0.22	-0.81	15,15,15,15	0
3	CL	P	301	1/1	0.97	0.21	-0.87	28,28,28,28	0
4	AMP	H	402	22/23	0.94	0.20	-0.89	72,75,78,78	0
3	CL	B	301	1/1	0.94	0.22	-0.92	25,25,25,25	0
4	AMP	J	401	23/23	0.89	0.21	-1.03	75,79,90,91	0
4	AMP	N	402	22/23	0.92	0.22	-1.09	81,87,90,90	0
4	AMP	X	402	22/23	0.91	0.23	-1.14	74,81,82,83	0
4	AMP	B	403	22/23	0.93	0.18	-1.26	80,83,84,85	0
3	CL	H	301	1/1	0.99	0.18	-1.31	23,23,23,23	0
4	AMP	J	402	22/23	0.92	0.19	-1.31	72,75,76,79	0
3	CL	R	301	1/1	0.97	0.15	-1.35	22,22,22,22	0
3	CL	X	301	1/1	0.98	0.19	-1.36	15,15,15,15	0
3	CL	D	301	1/1	0.95	0.20	-1.56	20,20,20,20	0
4	AMP	H	401	23/23	0.85	0.21	-1.56	78,85,89,90	0
3	CL	V	301	1/1	0.96	0.19	-1.58	13,13,13,13	0
4	AMP	V	402	22/23	0.93	0.18	-1.61	72,74,77,77	0
4	AMP	N	401	23/23	0.82	0.22	-1.62	82,82,88,88	0
4	AMP	H	403	22/23	0.93	0.17	-1.73	70,70,72,73	0
4	AMP	B	401	23/23	0.87	0.18	-1.92	79,80,86,86	0
4	AMP	J	403	22/23	0.93	0.18	-2.05	80,88,89,89	0
4	AMP	B	402	22/23	0.94	0.17	-2.35	77,85,86,86	0
3	CL	L	301	1/1	0.93	0.10	-2.78	44,44,44,44	0
4	AMP	X	401	23/23	0.92	0.20	-3.95	72,73,84,85	0
4	AMP	V	403	22/23	0.96	0.16	-	70,73,73,73	0

## 6.5 Other polymers

There are no such residues in this entry.