



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 20, 2023 – 01:52 AM EDT

PDB ID : 5K36  
Title : Structure of an eleven component nuclear RNA exosome complex bound to RNA  
Authors : Lima, C.D.; Zinder, J.C.  
Deposited on : 2016-05-19  
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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

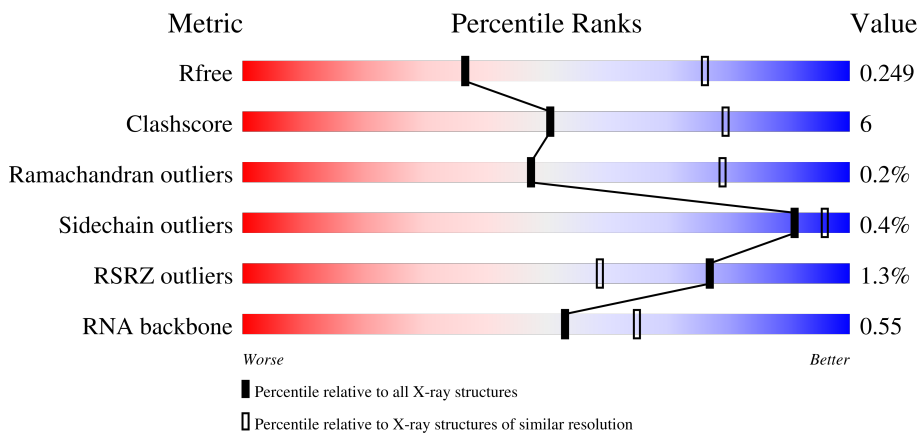
# 1 Overall quality at a glance

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))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	305	 85% 10% 5%
2	B	250	 80% 15% .
3	C	394	 72% 14% . 14%
4	D	225	 78% 20% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	269	<p>%</p> <p>79% 17%</p>
6	F	250	<p>2%</p> <p>71% 14% 14%</p>
7	G	244	<p>%</p> <p>83% 12% 5%</p>
8	H	363	<p>2%</p> <p>68% 12% 20%</p>
9	I	296	<p>7%</p> <p>65% 11% 24%</p>
10	J	559	<p>%</p> <p>71% 12% 16%</p>
11	K	1003	<p>80% 15% 5%</p>
12	L	17	<p>59% 29% 12%</p>
12	M	17	<p>24% 12% 6% 59%</p>

## 2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 30251 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 Exosome complex component RRP45.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	290	2255	1419	385	434	17	0	0	0

- Molecule 2 is a protein called Exosome complex component SKI6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	239	1883	1178	336	360	9	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP P46948
B	-2	PRO	-	expression tag	UNP P46948
B	-1	ASP	-	expression tag	UNP P46948
B	0	HIS	-	expression tag	UNP P46948

- Molecule 3 is a protein called Exosome complex component RRP43.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	340	2656	1683	446	516	11	0	0	0

- Molecule 4 is a protein called Exosome complex component RRP46.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	221	1691	1063	287	332	9	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP P53256
D	0	SER	-	expression tag	UNP P53256

- Molecule 5 is a protein called Exosome complex component RRP42.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	259	1997	1275	330	387	5	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-3	GLY	-	expression tag	UNP Q12277
E	-2	ASP	-	expression tag	UNP Q12277
E	-1	PRO	-	expression tag	UNP Q12277
E	0	HIS	-	expression tag	UNP Q12277

- Molecule 6 is a protein called Exosome complex component MTR3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	215	1651	1033	278	330	10	0	0	0

- Molecule 7 is a protein called Exosome complex component RRP40.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	232	1803	1152	295	345	11	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	GLY	-	expression tag	UNP Q08285
G	-2	ASP	-	expression tag	UNP Q08285
G	-1	PRO	-	expression tag	UNP Q08285
G	0	HIS	-	expression tag	UNP Q08285

- Molecule 8 is a protein called Exosome complex component RRP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	289	2245	1404	406	423	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	-3	GLY	-	expression tag	UNP P38792
H	-2	ASP	-	expression tag	UNP P38792
H	-1	PRO	-	expression tag	UNP P38792
H	0	HIS	-	expression tag	UNP P38792

- Molecule 9 is a protein called Exosome complex component CSL4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	224	1728	1079	305	337	7	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	-3	GLY	-	expression tag	UNP P53859
I	-2	ASP	-	expression tag	UNP P53859
I	-1	PRO	-	expression tag	UNP P53859
I	0	HIS	-	expression tag	UNP P53859

- Molecule 10 is a protein called Exosome complex exonuclease RRP6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	467	3829	2445	659	715	10	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	126	SER	-	expression tag	UNP Q12149
J	127	LEU	-	expression tag	UNP Q12149
J	128	MET	-	expression tag	UNP Q12149
J	238	ASN	ASP	engineered mutation	UNP Q12149

- Molecule 11 is a protein called Exosome complex exonuclease DIS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	955	7638	4829	1344	1430	35	0	0	0

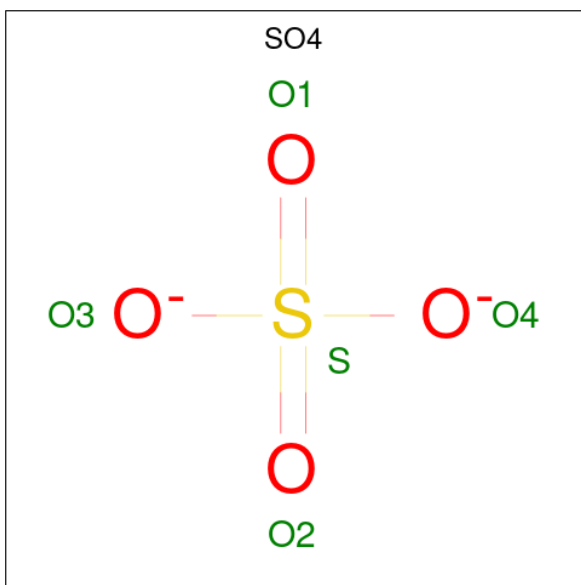
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-1	SER	-	expression tag	UNP Q08162
K	0	LEU	-	expression tag	UNP Q08162
K	171	ASN	ASP	engineered mutation	UNP Q08162
K	551	ASN	ASP	engineered mutation	UNP Q08162

- Molecule 12 is a RNA chain called RNA (17-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	M	7	Total	C	N	O	P	0	0	0
			148	67	26	48	7			
12	L	17	Total	C	N	O	P	0	0	0
			354	160	55	122	17			

- Molecule 13 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



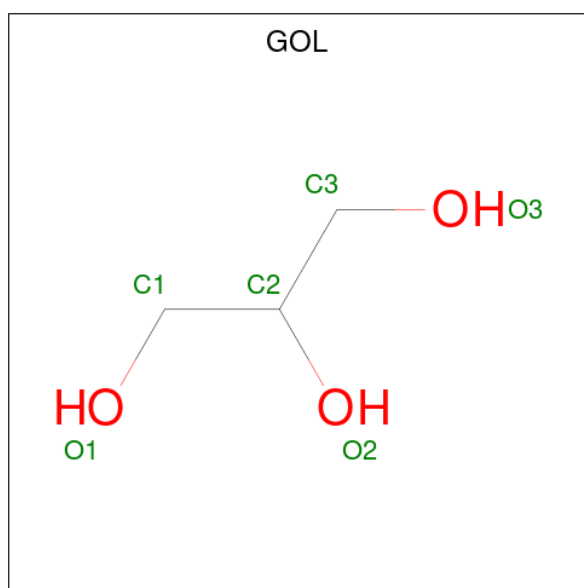
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	A	1	Total	O	S	0	0
			5	4	1		
13	B	1	Total	O	S	0	0
			5	4	1		
13	D	1	Total	O	S	0	0
			5	4	1		
13	E	1	Total	O	S	0	0
			5	4	1		
13	F	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	G	1	Total	O	S	0	0
			5	4	1		
13	I	1	Total	O	S	0	0
			5	4	1		
13	K	1	Total	O	S	0	0
			5	4	1		
13	K	1	Total	O	S	0	0
			5	4	1		
13	M	1	Total	O	S	0	0
			5	4	1		

- Molecule 14 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
14	A	1	Total	C	O	0	0
			6	3	3		
14	A	1	Total	C	O	0	0
			6	3	3		
14	A	1	Total	C	O	0	0
			6	3	3		
14	B	1	Total	C	O	0	0
			6	3	3		
14	H	1	Total	C	O	0	0
			6	3	3		
14	K	1	Total	C	O	0	0
			6	3	3		

Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
14	K	1	Total	C	O	0	0
			6	3	3		

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	K	1	Total	Zn	0	0
			1	1		

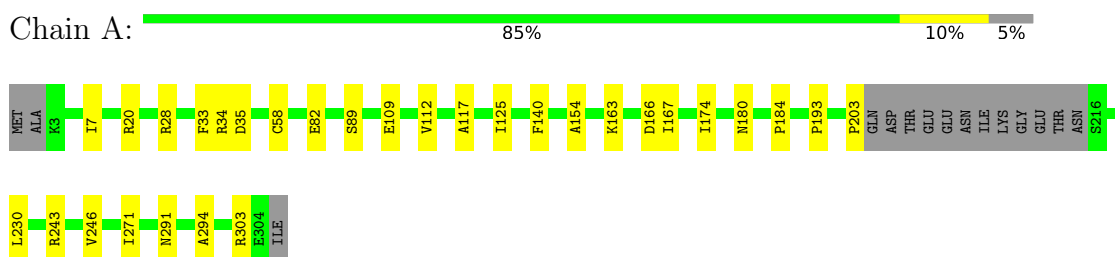
- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	52	Total	O	0	0
			52	52		
16	B	32	Total	O	0	0
			32	32		
16	C	18	Total	O	0	0
			18	18		
16	D	27	Total	O	0	0
			27	27		
16	E	8	Total	O	0	0
			8	8		
16	F	9	Total	O	0	0
			9	9		
16	G	30	Total	O	0	0
			30	30		
16	H	11	Total	O	0	0
			11	11		
16	I	4	Total	O	0	0
			4	4		
16	J	23	Total	O	0	0
			23	23		
16	K	64	Total	O	0	0
			64	64		
16	L	2	Total	O	0	0
			2	2		

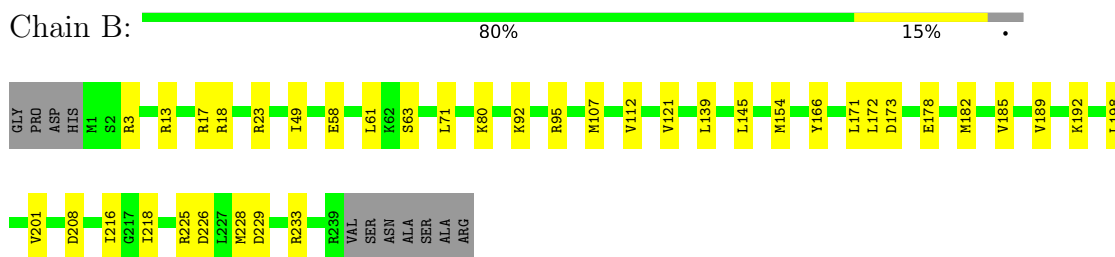
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

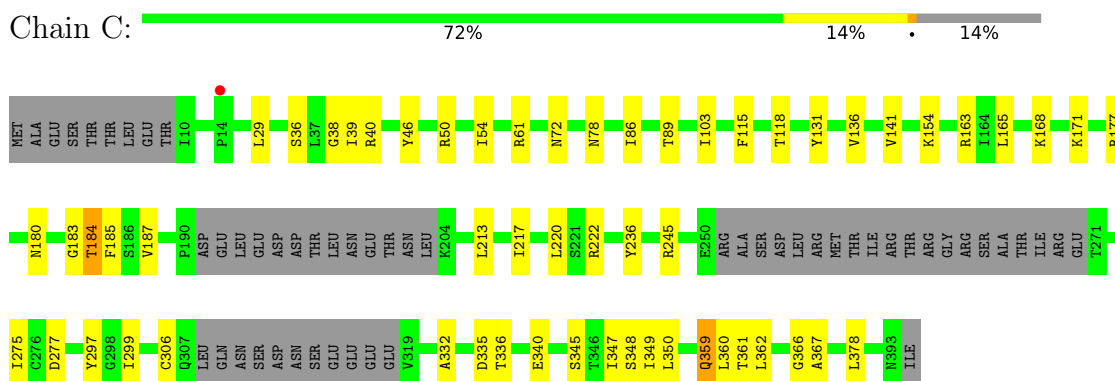
- Molecule 1: Exosome complex component RRP45



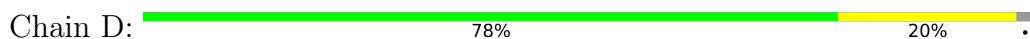
- Molecule 2: Exosome complex component SKI6

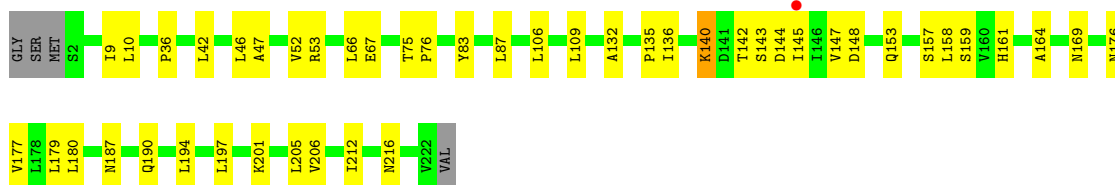


- Molecule 3: Exosome complex component RRP43

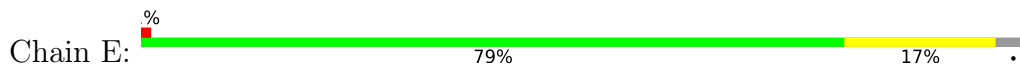


- Molecule 4: Exosome complex component RRP46

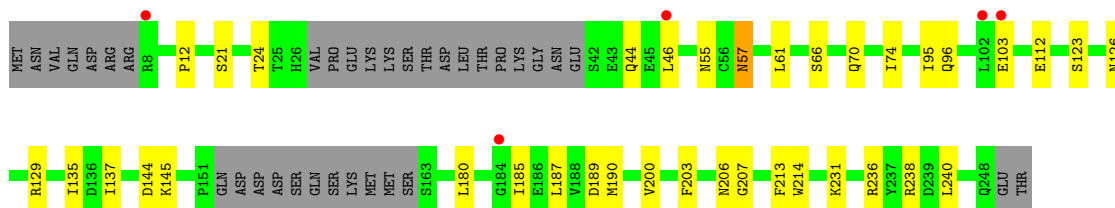




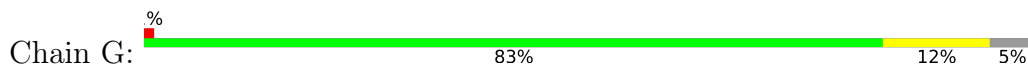
- Molecule 5: Exosome complex component RRP42



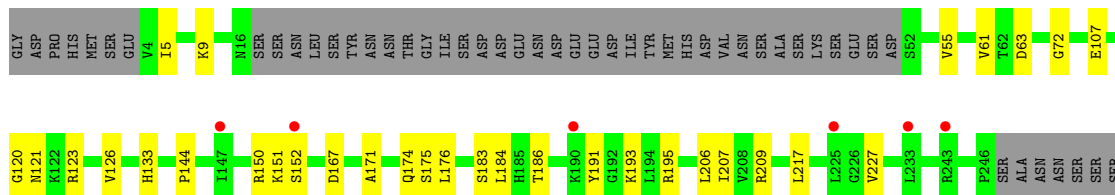
- Molecule 6: Exosome complex component MTR3



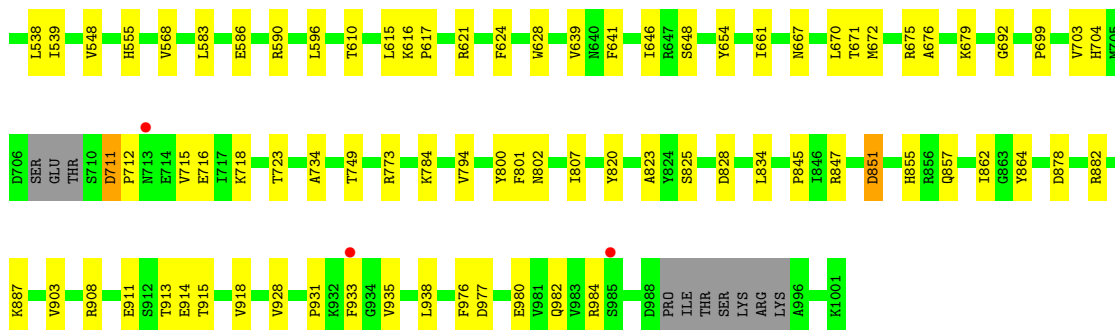
- Molecule 7: Exosome complex component RRP40



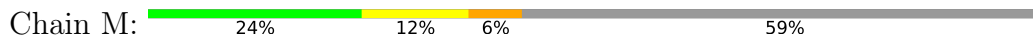
- Molecule 8: Exosome complex component RRP4







• Molecule 12: RNA (17-MER)



• Molecule 12: RNA (17-MER)



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	140.70Å 212.26Å 218.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	106.13 – 3.10 106.13 – 3.10	Depositor EDS
% Data completeness (in resolution range)	97.6 (106.13-3.10) 97.6 (106.13-3.10)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.40 (at 3.13Å)	Xtrriage
Refinement program	PHENIX 1.10_2155	Depositor
R, $R_{free}$	0.201 , 0.249 0.201 , 0.249	Depositor DCC
$R_{free}$ test set	5764 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.8	Xtrriage
Anisotropy	0.662	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 63.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.023 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	30251	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.20% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: SO4, GOL, ZN

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	0.23	0/2290	0.40	0/3088
2	B	0.23	0/1907	0.41	0/2566
3	C	0.23	0/2695	0.42	0/3647
4	D	0.23	0/1709	0.42	0/2319
5	E	0.24	0/2034	0.41	0/2764
6	F	0.24	0/1675	0.44	0/2263
7	G	0.24	0/1839	0.41	0/2493
8	H	0.24	0/2280	0.42	0/3078
9	I	0.24	0/1751	0.43	0/2366
10	J	0.23	0/3914	0.39	0/5310
11	K	0.23	0/7787	0.40	0/10550
12	L	0.21	0/394	0.92	3/609 (0.5%)
12	M	0.13	0/165	0.65	0/254
All	All	0.23	0/30440	0.42	3/41307 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
12	L	12	U	C2-N1-C1'	7.33	126.49	117.70
12	L	12	U	N1-C2-O2	6.95	127.67	122.80
12	L	12	U	N3-C2-O2	-6.43	117.70	122.20

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	2255	0	2251	20	0
2	B	1883	0	1940	26	0
3	C	2656	0	2705	39	0
4	D	1691	0	1745	27	0
5	E	1997	0	2037	31	0
6	F	1651	0	1624	27	0
7	G	1803	0	1794	18	0
8	H	2245	0	2265	27	0
9	I	1728	0	1729	22	0
10	J	3829	0	3839	46	0
11	K	7638	0	7678	89	0
12	L	354	0	178	4	0
12	M	148	0	75	3	0
13	A	5	0	0	0	0
13	B	5	0	0	0	0
13	D	5	0	0	0	0
13	E	5	0	0	0	0
13	F	5	0	0	0	0
13	G	5	0	0	1	0
13	I	5	0	0	0	0
13	K	10	0	0	0	0
13	M	5	0	0	0	0
14	A	18	0	24	2	0
14	B	6	0	8	0	0
14	H	6	0	8	1	0
14	K	12	0	16	3	0
15	K	1	0	0	0	0
16	A	52	0	0	1	0
16	B	32	0	0	1	0
16	C	18	0	0	0	0
16	D	27	0	0	0	0
16	E	8	0	0	0	0
16	F	9	0	0	1	0
16	G	30	0	0	0	0
16	H	11	0	0	0	0
16	I	4	0	0	0	0
16	J	23	0	0	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	K	64	0	0	0	0
16	L	2	0	0	0	0
All	All	30251	0	29916	343	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 6.

The worst 5 of 343 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)
3:C:362:LEU:HB3	4:D:180:LEU:HB3	1.69	0.74
9:I:51:ARG:HB3	10:J:548:ARG:HH22	1.56	0.70
8:H:144:PRO:HB3	8:H:278:TRP:HB3	1.73	0.69
11:K:878:ASP:OD2	11:K:882:ARG:NH1	2.27	0.67
5:E:215:ILE:HA	5:E:226:ILE:HG22	1.77	0.65

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	286/305 (94%)	277 (97%)	9 (3%)	0	100	100
2	B	237/250 (95%)	226 (95%)	11 (5%)	0	100	100
3	C	332/394 (84%)	306 (92%)	25 (8%)	1 (0%)	41	73
4	D	219/225 (97%)	211 (96%)	7 (3%)	1 (0%)	29	64
5	E	255/269 (95%)	240 (94%)	14 (6%)	1 (0%)	34	69
6	F	209/250 (84%)	190 (91%)	19 (9%)	0	100	100
7	G	228/244 (93%)	216 (95%)	12 (5%)	0	100	100
8	H	283/363 (78%)	266 (94%)	16 (6%)	1 (0%)	34	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	I	216/296 (73%)	201 (93%)	15 (7%)	0	100	100
10	J	459/559 (82%)	438 (95%)	20 (4%)	1 (0%)	47	79
11	K	945/1003 (94%)	900 (95%)	44 (5%)	1 (0%)	51	83
All	All	3669/4158 (88%)	3471 (95%)	192 (5%)	6 (0%)	47	79

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	184	THR
11	K	711	ASP
4	D	143	SER
8	H	152	SER
10	J	539	VAL

### 5.3.2 Protein sidechains [i](#)

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	252/266 (95%)	251 (100%)	1 (0%)	91	96
2	B	213/221 (96%)	213 (100%)	0	100	100
3	C	299/349 (86%)	297 (99%)	2 (1%)	84	93
4	D	195/198 (98%)	194 (100%)	1 (0%)	88	94
5	E	235/243 (97%)	234 (100%)	1 (0%)	91	96
6	F	185/219 (84%)	182 (98%)	3 (2%)	62	84
7	G	202/212 (95%)	202 (100%)	0	100	100
8	H	248/314 (79%)	247 (100%)	1 (0%)	91	96
9	I	187/243 (77%)	187 (100%)	0	100	100
10	J	431/516 (84%)	429 (100%)	2 (0%)	88	94
11	K	857/903 (95%)	855 (100%)	2 (0%)	93	97
All	All	3304/3684 (90%)	3291 (100%)	13 (0%)	91	96

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	129	ARG
8	H	209	ARG
11	K	851	ASP
10	J	409	LEU
11	K	624	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
10	J	395	ASN
11	K	275	GLN
11	K	21	GLN
11	K	279	GLN
5	E	173	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
12	L	16/17 (94%)	4 (25%)	0
12	M	6/17 (35%)	2 (33%)	0
All	All	22/34 (64%)	6 (27%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
12	M	12	U
12	M	17	A
12	L	4	U
12	L	8	U
12	L	10	U

There are no RNA pucker outliers to report.

## 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 1 is monoatomic - leaving 17 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
14	GOL	K	2005	-	5,5,5	0.37	0	5,5,5	0.20	0
13	SO4	B	301	-	4,4,4	0.14	0	6,6,6	0.04	0
13	SO4	K	2002	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	D	301	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	F	301	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	K	2003	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	M	101	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	G	301	-	4,4,4	0.14	0	6,6,6	0.05	0
14	GOL	A	402	-	5,5,5	0.36	0	5,5,5	0.26	0
14	GOL	B	302	-	5,5,5	0.37	0	5,5,5	0.29	0
13	SO4	A	401	-	4,4,4	0.14	0	6,6,6	0.05	0
14	GOL	K	2004	-	5,5,5	0.38	0	5,5,5	0.25	0
14	GOL	A	403	-	5,5,5	0.38	0	5,5,5	0.26	0
13	SO4	I	301	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	E	301	-	4,4,4	0.15	0	6,6,6	0.05	0
14	GOL	H	401	-	5,5,5	0.37	0	5,5,5	0.25	0
14	GOL	A	404	-	5,5,5	0.36	0	5,5,5	0.27	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
14	GOL	K	2005	-	-	2/4/4/4	-
14	GOL	B	302	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
14	GOL	A	402	-	-	2/4/4/4	-
14	GOL	K	2004	-	-	2/4/4/4	-
14	GOL	A	403	-	-	2/4/4/4	-
14	GOL	H	401	-	-	2/4/4/4	-
14	GOL	A	404	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
14	H	401	GOL	O1-C1-C2-C3
14	K	2004	GOL	O1-C1-C2-C3
14	K	2005	GOL	O1-C1-C2-C3
14	H	401	GOL	O1-C1-C2-O2
14	K	2005	GOL	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	K	2005	GOL	2	0
13	G	301	SO4	1	0
14	A	402	GOL	1	0
14	K	2004	GOL	1	0
14	A	403	GOL	1	0
14	H	401	GOL	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	290/305 (95%)	-0.17	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	28, 47, 83, 114	0
2	B	239/250 (95%)	-0.01	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	25, 50, 102, 137	0
3	C	340/394 (86%)	0.07	1 (0%) <span style="border: 1px solid blue; padding: 2px;">94</span> <span style="border: 1px solid blue; padding: 2px;">88</span>	50, 83, 127, 173	0
4	D	221/225 (98%)	-0.11	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">81</span>	36, 59, 87, 121	0
5	E	259/269 (96%)	-0.01	3 (1%) <span style="border: 1px solid blue; padding: 2px;">79</span> <span style="border: 1px solid blue; padding: 2px;">61</span>	43, 91, 125, 145	0
6	F	215/250 (86%)	0.27	5 (2%) <span style="border: 1px solid blue; padding: 2px;">60</span> <span style="border: 1px solid red; padding: 2px;">39</span>	62, 99, 144, 162	0
7	G	232/244 (95%)	-0.03	2 (0%) <span style="border: 1px solid blue; padding: 2px;">84</span> <span style="border: 1px solid blue; padding: 2px;">69</span>	36, 62, 119, 157	0
8	H	289/363 (79%)	0.20	8 (2%) <span style="border: 1px solid blue; padding: 2px;">53</span> <span style="border: 1px solid red; padding: 2px;">30</span>	48, 104, 147, 181	0
9	I	224/296 (75%)	0.76	22 (9%) <span style="border: 1px solid red; padding: 2px;">7</span> <span style="border: 1px solid red; padding: 2px;">2</span>	55, 114, 171, 194	0
10	J	467/559 (83%)	-0.00	4 (0%) <span style="border: 1px solid blue; padding: 2px;">84</span> <span style="border: 1px solid blue; padding: 2px;">69</span>	50, 84, 160, 183	0
11	K	955/1003 (95%)	-0.02	4 (0%) <span style="border: 1px solid blue; padding: 2px;">92</span> <span style="border: 1px solid blue; padding: 2px;">84</span>	30, 65, 124, 163	0
12	L	17/17 (100%)	0.52	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	65, 128, 149, 166	0
12	M	7/17 (41%)	-0.08	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	61, 63, 86, 116	0
All	All	3755/4192 (89%)	0.06	50 (1%) <span style="border: 1px solid blue; padding: 2px;">77</span> <span style="border: 1px solid blue; padding: 2px;">59</span>	25, 75, 141, 194	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	I	25	ASN	4.1
9	I	109	SER	4.0
9	I	106	ILE	3.6
6	F	103	GLU	3.5
9	I	21	TYR	3.2

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
13	SO4	E	301	5/5	0.78	0.12	124,141,148,155	0
13	SO4	G	301	5/5	0.85	0.46	105,105,132,150	0
13	SO4	K	2002	5/5	0.85	0.14	98,104,145,151	0
14	GOL	A	404	6/6	0.86	0.36	81,85,87,92	0
14	GOL	B	302	6/6	0.88	0.40	68,72,79,90	0
13	SO4	F	301	5/5	0.89	0.19	107,147,152,155	0
13	SO4	I	301	5/5	0.90	0.20	109,120,154,155	0
14	GOL	A	403	6/6	0.90	0.28	56,66,78,80	0
13	SO4	M	101	5/5	0.91	0.22	95,102,122,122	0
13	SO4	K	2003	5/5	0.92	0.19	84,88,113,113	0
14	GOL	H	401	6/6	0.92	0.32	55,75,90,92	0
14	GOL	K	2004	6/6	0.92	0.34	45,59,75,76	0
13	SO4	D	301	5/5	0.94	0.17	82,86,114,127	0
14	GOL	K	2005	6/6	0.94	0.31	58,66,83,90	0
13	SO4	B	301	5/5	0.95	0.14	98,99,127,134	0
14	GOL	A	402	6/6	0.95	0.35	57,63,81,87	0
13	SO4	A	401	5/5	0.99	0.19	45,54,64,66	0
15	ZN	K	2001	1/1	1.00	0.20	44,44,44,44	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.